The Dynamic Behaviour of Road Traffic Flow : Stability or Chaos ?

A thesis submitted to Middlesex University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

XIAOYAN ZHANG

School of Mathematics and Statistics
Middlesex University

December 1995

To my family

ACKNOWLEDGEMENTS

My thanks first of all go to the Director of Studies and my first supervisor, Mr. D. F. Jarrett, for his good supervision and many helpful discussions throughout the research, for his thorough and thoughtful reading of the thesis, and for the clear and concise comments he made to the thesis.

Thanks are gratefully extended to my second supervisor, Professor C. C. Wright, for his helpful suggestions and comments about the research and the thesis, and for his encouragements.

I would like to acknowledge Middlesex University for providing a three-year studentship to support the research in this thesis.

The librarians in the Library of Middlesex University at Hendon gave me many helps in getting materials needed for the research. The calculations in the research are made on the VAX computer system in Middlesex University. The diagrams in the thesis are produced by MATLAB (The MathWorks Inc., 1993), also on the VAX computer system in Middlesex University. The thesis is typed by T^3 Scientific Word Processing System (TCI Software Research, 1985).

I take full responsibility for any errors or omissions in this thesis.

ABSTRACT

The objective of this thesis is to investigate the dynamic behaviour of road traffic flow based on theoretical traffic models. Three traffic models are examined: the classical car-following model which describes the variations of speeds of cars and distances between the cars on a road link, the logit-based trip assignment model which describes the variations of traffic flows on road links in a road network, and the dynamic gravity trip distribution model which describes the variations of flows between O–D pairs in an O–D network.

Some dynamic analyses have been made of the car-following model in the literature (Chandler *et al.*, 1958, Herman *et al.*, 1959, Disbro & Frame, 1990, and Kirby and Smith, 1991). The dynamic gravity model and the logit-based trip assignment model are both suggested by Dendrinos and Sonis (1990) without detailed analysis. There is virtually no previous dynamic analysis of trip distribution, although there are some dynamic considerations of trip assignment based on other assignment models (Smith, 1984 and Horowitz, 1984).

In this thesis, the three traffic models are considered as dynamical systems. The variations of traffic characteristics are investigated in the context of nonlinear dynamics. Equilibria and oscillatory behaviour are found in all three traffic models; complicated behaviour including period doubling and chaos is found in the gravity model. Values of parameters for different types of behaviour in the models are given. Conditions for the stability of equilibria in the models are established. The stability analysis of the equilibrium in the car-following model is more general here than that in the literature (Chandler *et al.*, 1958, Herman *et al.*, 1959). Chaotic attractors found in the gravity model are characterized by Liapunov exponents and fractal dimension.

The research in this thesis aims at understanding and predicting traffic behaviour under various conditions. Traffic systems may be monitored, based on these results, to achieve a stable equilibrium and to avoid instabilities and chaos.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1. INTRODUCTION

1.1. BACKGROUND

This thesis is concerned with the dynamic behaviour of road traffic flow. We all know that traffic characteristics such as speed and flow vary with time, but do we know how they vary?

In order to understand road traffic behaviour, analysts have developed traffic models, in the form of mathematical equations in most cases, to replicate real traffic systems. Although these models are simplified representations of real systems, they may be very complicated from mathematical point of view, and often need to be tested by large amounts of data. The process of traffic modelling involves model development, model analysis, and model calibration and validation. Through this process we can learn a lot about both the properties of the model and the real traffic system. Traffic can thus be managed, controlled, and planned more sensibly.

Traffic systems consist of roads, moving vehicles or traffic flows, and people (drivers and passengers). In traffic modelling, we may either concentrate on traffic flows on a stretch of road only or, more generally, look at the flows in a road network. Consider first traffic flow on a road link between junctions. Cars which move close enough may interact with each other. A driver receives visual information about the motions of neighbouring cars, particularly the car immediately in front. He or she makes judgements about the positions and speeds of his or her car and the car in front, and responds accordingly, but not necessarily immediately. There may be a short time for reaction. Drivers may accelerate or decelerate if the car in front does so. This system is modelled by the *car-following model*. In its simplest, linear form, the car-following model states that the acceleration of a car at time t is proportional to its speed relative to the car in front at time $t-\tau$, where τ is the *delay* or the *reaction time*. More general car-following models are nonlinear and assume that the acceleration of a car depends not only on the relative speed to the car in front, but also on their

distance apart and its own speed.

The traffic flow on a road link has also been modelled from a different perspective. Instead of treating each car individually like one does in the car-following model, traffic is considered as a continuum and is characterized by aggregate variables such as the speed, flow, and density. The *fluid model* describes the relationship of these variables and their variations with time and space.

The car-following model and the fluid model deal with the traffic flow on a stretch of road. There are other traffic models concerned with the allocation or the distribution of traffic flows in a road network or an area. Traffic flows are often referred to as numbers of trips in these models.

In a road network, each driver is making a *trip*, or a journey from an *origin*, the place where a journey starts, to a *destination*, the place where the journey ends. Drivers often have more than one possible route (a chain of road links) to choose between a particular origin-destination pair (O–D pair). They naturally choose the best route or what they think is the best route. The main concern in route choice is travel time or some more general notion of travel cost. A cheaper route may become more crowded and so more costly to travel. In the meantime, different drivers can have different beliefs about the best route. As a result, different drivers travelling between the same two places often use different routes; the flow between each O–D pair is shared among the routes connecting the origin and the destination. The *trip assignment model* describes the ways in which the traffic flow between each O–D pair is allocated to the routes according to their travel costs. A *user equilibrium* is achieved when the flow on each route is such that no road user can improve his or her travel cost by changing routes.

Not only can one choose which route to use in a road network, but also one can choose where to live and where to work, to do shopping, to travel, etc. The choices are again often made based on some kind of cost such as the distance, or the travel time. Consider a large area composed of several origins and destinations. The *trip generation model* determines the total number of trips generated from each origin and attracted to each destination from factors such as land use and socio-economic conditions in the area. Given the total number of trips from each origin and to each destination, the *trip distribution model* will determine the number of trips between each O–D pair, possibly according to the

travel costs between the two places. The best known trip distribution model is the *gravity model*, which originated from an analogy with Newton's gravitational law. It is assumed that the number of trips between an origin and a destination is proportional to the number of trips from the origin and the number of trips to the destination, and is inversely proportional to some measure of cost such as the square of the distance between the origin and the destination.

Traffic models such as those mentioned above describe the relationships among traffic variables, and related factors such as travel costs. In reality, these traffic characteristics (the flow, speed, and density in traffic models at the road link level; and the number of trips on each route, or between each O–D pair at the network level) vary with time. But not all traffic models take this fact into account.

Traffic models can be divided into two categories, static and dynamic, according to whether the time variation is included explicitly or not. Static models consider only one state of traffic which is implicitly assumed to prevail in a traffic system. The static trip assignment model is a typical example. Many trip assignment models have been developed so that their solutions satisfy the user equilibrium conditions. The user equilibrium is often formulated as the solution of a mathematical programming problem. Traffic flows are allocated to each route such that the travel cost of each driver is minimized. However, it may be doubtful whether traffic flows in a road network will stick to the equilibrium, although the underlying assumption that drivers try to minimize their own travel cost is natural and plausible. If the system is sensitive to various disturbances, which are inevitable in practice, then the equilibrium can hardly last in the system even if it exists. In other words, the equilibrium may not be *stable*. To seek an unstable equilibrium is just like trying to stand an egg on its end.

Dynamic models describe the evolution of the system modelled. To check the stability of an equilibrium, we only need to simulate the model, starting somewhere near the equilibrium as if the system is perturbed a little from the equilibrium, and see if the system will stay nearby or even go back to the equilibrium eventually.

The motion of traffic flows on a road link or a road network can be considered as a dynamical system. Dynamic models of traffic flow describe the time evolution of such variables as flow rate (traffic flow), and speed, which characterize the

state of the system. The time evolution of a dynamical system is normally modelled by (deterministic) differential or difference equations which relate the rate of change of the state to the current values of the state variables and necessary parameters. In practical applications, these equations will usually be nonlinear. A solution of the equations gives one history of the system, which traces out a *trajectory* or an *orbit* in the *phase space* of the system, or the (multidimensional) space of the variables of the system. Each point of the trajectory represents a state of the system at a certain time.

Given an initial condition or a starting point, many dynamical systems exhibit a start-up transient, after which the motion settles down towards some form of long-term recurrent behaviour. Motions from neighboring initial values tend to converge towards the stable attracting solutions called *attractors*. There are basically four types of attractor: *point attractors*, which is a stable equilibrium where the variables of the system are constants, *periodic attractors* on which the state of the system varies periodically, *quasi-periodic attractors* on which the state of the system varies regularly but does not repeat itself exactly, and *chaotic attractors* on which the state of the system varies irregularly. Equilibria and periodic motions are familiar to us; chaotic motions are not fully understood yet.

The car-following model mentioned above is an example of a dynamical system. It is a system of delay-differential equations describing the motion of a line of cars on a road link. The solution of the equations describes the speed of each car and the spacing between the cars as a function of time. It may settle down with time to an equilibrium where the speeds and spacings are constants. Similarly, a dynamic trip assignment model can be used to describe the (daily, weekly, etc.) variations of traffic flows in a road network. It would be very useful if the flow pattern in a dynamic trip assignment model approached the user equilibrium. However, an equilibrium in a dynamic trip assignment model may or may not be the same as the user equilibrium, depending on how the system is modelled. If it is assumed that a driver would rather choose the route which may be longer but with a more scenic view, then an equilibrium in the model (if it exists) may not guarantee that all drivers travel at the minimum cost. The stability of an equilibrium is most important in traffic analysis. An unstable equilibrium is practically useless. Unless starting *exactly* at the equilibrium, a trajectory rarely heads for an unstable equilibrium.

There has been a great interest in nonlinear dynamics since the 1960s, because of

the wide applications of powerful computers and rapid progress in geometric and topological methods in dynamics. Chaotic behaviour is a revolutionary discovery in deterministic systems. There is no widely accepted definition of a chaotic attractor. However, it may be more helpful to know how a chaotic system behaves than to define it.

Although trajectories near a chaotic attractor converge to the attractor, two nearby trajectories on the attractor diverge *exponentially* fast with time. In other words, the system is extremely *sensitive to initial conditions*. It is this feature which distinguishes chaotic and non-chaotic behaviour. Practically we cannot deal with infinitely precise numbers; we normally have to predict the future of the system from imprecise initial conditions. The orbital divergence makes it impossible to tell conclusively what is going to happen next. The behaviour appears to be *stochastic*, yet in a *deterministic* system. With a deterministic system, such as motions of bodies described by Newtonian mechanics, it had long been believed that the future is uniquely determined from the past and that a small error remains small for all time. Today, this cannot be taken for granted any longer. In a chaotic system, an initially insignificant error, *however small it may be*, will eventually become significant and intolerable.

Chaotic behaviour is irregular, unpredictable, and complicated. The unpredictability can be measured by *Liapunov exponents*, which tells how fast neighbouring trajectories diverge or converge exponentially. The exponential divergence cannot go too far: trajectories will wind back on themselves because the size of an attractor is finite. This *stretching and folding* makes a chaotic attractor very complicated such that it normally has a non-integer or *fractal* dimension.

It may be surprising to know that many practical systems which seem to be stochastic can be modelled by deterministic models, such as the weather forecast models. Chaotic behaviour has been found in some of these systems (for example, Lorenz, 1963). Even a very simple mathematical equation can give rise to exceptionally complicated solutions. Although traffic phenomena seem stochastic, many traffic models are deterministic, partly because of their relative simplicity. A stochastic model can be very difficult to handle and to apply if they are made reasonably realistic. Many stochastic models have to be approximated by deterministic ones so that they are approachable. Queueing models, for example, are stochastic. They model a system in which customers

arrive, queue, and are served. They have been applied to model road traffic systems. Newell (1982) has observed, in his work on applications of queueing theory to traffic systems, that "deterministic approximations have found application to a much wider range of practical problems than the stochastic theory simply because the stochastic analysis of even the simplest systems which involve several servers or customer types is too tedious to be of much practical value". However, a deterministic model can give rise to stochastic behaviour, as we have seen. One would not be too surprised if chaos were found in (deterministic) traffic models. That would mean that apparently stochastic traffic phenomena could turn out to have a deterministic mechanism. In fact, it is almost impossible to determine if a given random appearing behaviour is probabilistic or deterministic. The discovery of chaos has made the distinction between deterministic and stochastic behaviour much more blurred.

The steady state of a dynamical system may depend on the initial state and also on other control conditions, which are normally represented by one or more parameters in the model. As these conditions change, the long term behaviour of the system may change, either quantitatively, or qualitatively. At a certain value of a parameter, an equilibrium may *bifurcate* into a periodic motion, or, a periodic orbit may become irregular or chaotic. If an equilibrium of a dynamical system occurs only under very strict conditions, represented by a narrow range of values of the parameters, then it may not be easy to maintain the equilibrium in the system.

So far as road traffic flow is concerned, equilibrium and stability are always desired. An attracting equilibrium is highly desirable to road users, planners, and traffic engineers alike. In traffic studies, many traffic models are formulated and solved to obtain a user equilibrium without any stability analysis being carried out. What is more, these models are then used to give guidance to drivers in practice, aiming to get a smooth, efficient traffic flow. In real traffic, however, it is obvious that an equilibrium may not always be achieved. Instabilities inevitably exist, though their effects may be very weak. In dense traffic, where drivers follow each other very closely, small disturbances like the acceleration or deceleration of one vehicle might be preserved or amplified along the line of vehicles or over time. This sensitive dependence on initial conditions implies instability. These phenomena may raise problems in traffic management, and may even result in accidents when serious.

In traffic systems, initial conditions and values of parameters can never be determined with perfect precision. It is therefore necessary to treat traffic as a dynamical system and to examine the behaviour of the system under various circumstances.

Computer technology has been developed in the form of artificial intelligence systems in cars (Bender and Fenton, 1970, and Kapur, 1971) and in automated route guidance systems (Watling and Van Vuren, 1993). The development of these control systems frequently suffers from the lack of dynamic modelling and analysis of traffic systems. Watling and Van Vuren (1993) quoted Smith and Ghali's comment (Smith and Ghali, 1991) to describe the situation: "technological developments are taking place at a much greater rate than improvements in the understanding and computer modelling of networks". One begins to wonder what would happen if the underlying traffic models in these control systems had instabilities or chaos built in.

1.2. THE AIM

The aim of this thesis is to investigate the dynamic behaviour of road traffic flow, based on deterministic, dynamic traffic models. It is intended to find out how traffic variables change with time under a governing equation and how the initial conditions and parameter values affect these changes. This knowledge is very important in both theoretical studies and applications.

Traffic flows are described at different levels of scope: the road link level, the road network level, and the origin-destination network level. Accordingly, it will be attempted to achieve the following three specific objectives:

(1) To investigate the variations of the speeds of the cars and the distances between the cars on a stretch of road.

(2) To investigate the variations of traffic flows on the routes and links in a road network.

(3) To investigate the variations of traffic flows between O–D pairs in an O–D network.

One dynamic traffic model in each of these three areas will be selected for the purpose of dynamic analysis. The models will then be examined to identify all possible long-term behaviour of traffic flows under various conditions. In particular, the different possible kinds of attractor will be found and described. Conditions for the stability of equilibrium will be established. Where an attractor appears to be chaotic, the chaos will be established and quantified by calculating Liapunov exponents and fractal dimensions.

One of the objectives of traffic modelling is to explain and to interpret traffic flow phenomena. By analyzing dynamic traffic models, the behaviour of traffic systems can be better understood. Dynamic analysis is also helpful in improving traffic models so that they are more realistic. The ultimate aim of traffic science is to achieve better traffic conditions. If it is found that there is more than one equilibrium in a system, which is not uncommon, then traffic can be monitored to achieve the desired equilibrium. If, on the other hand, unwanted behaviour such as instability or chaos is found for some values of parameters, then it may be possible to avoid it.

1.3. A BRIEF LITERATURE REVIEW

There has been much previous research on traffic dynamics. Many of them have focused on the stability of equilibrium solutions. On the other hand, the developments of chaos theory have stimulated traffic modellers. Some attempts have been made to identify chaos in traffic. Here, a brief review of the literature is given; a more detailed review can be found in Chapter 2.

First, there are two different models for describing the motion of vehicles along a road link: the car-following model and the fluid model. Most dynamic considerations of traffic flow on a road link are based on the car-following model. Chandler *et al.* (1958) and Herman *et al.* (1959) investigated the stability of the linear car-following model. Unwin *et al.* (1967) examined a special nonlinear car-following model, namely, the reciprocal model. Only two cars, a leader and a follower, were considered. More recently, two attempts have been made to identify possible chaotic behaviour in the car-following model. Disbro and Frame (1990) claimed to have found chaos in the model while Kirby and Smith (1991) found no evidence of chaos in the car-following model. This disagreement about the existence of chaos in the car-following model, together with the fact that only

the linear car-following model has been considered for stability (except Unwin *et al.*, 1967), motivated the study of the model in this thesis.

The fluid model is a system of partial differential equations. The dynamic behaviour of this model has been investigated by Kühne (1991) and Kühne and Beckschulte (1993). An apparently irregular behaviour was found; the authors did not conclude if the behaviour is chaotic or not. Dynamic study of partial differential equations is difficult both theoretically and numerically. Consequently, many dynamic models of partial differential equations are reduced to ordinary differential equations for the analysis.

Secondly, for traffic flows at the network level, there are the trip distribution model and the trip assignment model. Most traffic models at this level are static. Dendrinos and Sonis (1990) proposed a dynamic gravity model, with its potential for showing chaos, but they made no detailed study of the model. This dynamic gravity model is another model to be examined in this thesis.

As for the trip assignment model, two stability analyses are relevant here. Horowitz (1984) proposed a discrete-time assignment model for a simple road network of one O–D pair connected by two road links. The condition for the stability of the equilibrium in the model was provided. Smith (1984), on the other hand, suggested an assignment model consisting of a system of ordinary differential equations. The equilibrium of this dynamic system coincides with the user equilibrium, which Smith proved to be stable. One of the widely used trip assignment models is the logit choice model, first suggested by Dial (1971). This model is static. Dendrinos and Sonis (1990) proposed a dynamic logit choice model, which will be explored here for dynamic considerations in trip assignment.

1.4. MAIN CONTENTS OF THE THESIS

The dynamic behaviour of traffic flows at a road link, in a road network, and in an O–D network will be investigated based on, respectively, the car-following model, the dynamic gravity model, and the dynamic logit-based trip assignment model. The car-following model is a classical model while the dynamic gravity model and the logit assignment model are suggested by Dendrinos and Sonis (1990), as mentioned above.

The dynamic models of traffic flow will be analysed in the context of nonlinear dynamics. The time variation of characteristics of traffic flow is given by solutions of dynamic traffic models in the form of differential or difference equations. An *equilibrium* solution is one where these variables are constant over time. An equilibrium is said to be *stable* if the solutions converge to the equilibrium from any initial conditions sufficiently close to the equilibrium. As the value of a parameter in the model changes, the equilibrium may bifurcate to an oscillation or eventually to chaos.

In general, dynamic models of nonlinear differential or difference equations cannot be solved analytically; theoretical analysis can be made only to a limited extent. Numerical integration or iteration, together with geometrical methods, are very useful here. The idea is to visualize the states of the system as points in the phase space, instead of dealing with symbols and formulae. Thus, the *phase portrait* of an equilibrium is a single point in the space; a periodic solution is a closed loop; a chaotic attractor is a complicated object full of irregular trajectories, which may be beautiful! Although it is not possible to view a space of a dimension higher than three, it can be projected to a two dimensional space so that two variables can be examined. Different variables may be selected and viewed.

Equilibria are most important in traffic; their existence, uniqueness, and stability will be examined theoretically. More complicated behaviour will be examined by geometric methods such as time series plots, phase portraits projections based on numerical solutions. Sometimes it may be difficult to tell if a solution is regular or chaotic simply by viewing pictures of it. Spectral analysis may be helpful to differentiate periodic and non-periodic behaviour. If the behaviour appears to be chaotic, it will be examined further by calculating Liapunov exponents and fractal dimensions.

This thesis consists of seven chapters. A general introduction has been given here in the first chapter. Traffic models are described and reviewed in Chapter 2, where all relevant traffic models are described and suitable ones selected. A brief introduction of nonlinear dynamics is given in Chapter 3. Descriptions in these two chapters are general; more details will be given in due course in the later chapters. The three traffic models are examined in turn in the subsequent three chapters. Finally, the conclusions are summarized in the last chapter.

CHAPTER 2. DYNAMIC MODELS OF ROAD TRAFFIC FLOW

The main objective of this research is to investigate the dynamic behaviour of theoretical models of traffic flow. In this chapter, traffic models concerned with traffic dynamics are described and reviewed. Particular models are chosen for the research; some models which are not suitable here are modified for the purpose of this research. A general outline of traffic modelling is given in the first section. Particular models of interest are discussed in the subsequent sections.

2.1. TRAFFIC MODELS

The motion of road traffic is often characterized by such variables as displacement, speed, acceleration, traffic density, traffic volume, and so on. These variables will be explained when they are met. Traffic models describe relationships among these variables and how these variables vary over time and space. The purpose of mathematical modelling of road traffic flow is to interpret and to understand traffic behaviour so that appropriate actions can be taken to make traffic operate efficiently, safely, and economically. In the short term, these actions might consist of traffic management and control measures. For example, traffic signals can be timed so that traffic can move more smoothly; and instructions and information can be provided to drivers so that they can drive safely and quickly. In the longer term, congestion can be avoided in advance by traffic planning. This aims to forecast traffic changes (normally as a result of the development of land use and socio-economic activities) and traffic flow distributions over an area or a road network; the planning of a future stretch of road or road network can be based on these forecasts.

Like all other mathematical models, traffic models can be divided into deterministic models and stochastic models according to whether stochastic effects are considered. Traffic behaviour seems inherently stochastic, but many

traffic models are deterministic. One advantage of deterministic models over stochastic ones is that they are easier to deal with, as has been mentioned in Chapter 1. Traffic models can also be divided into static models and dynamic models according to whether time variations are involved explicitly. It is necessary to point out that many traffic models are static, despite the fact that traffic characteristics are bound to change with time. To investigate traffic dynamics a dynamic model is essential. Therefore, it is deterministic and dynamic traffic models which are of interest here, more specifically, traffic models in the form of differential or difference equations, which govern the time variation of traffic characteristics.

Traffic flows can be described either at a link level or at a network level. At a link level, traffic flow on only one road link between junctions or even on a single lane is considered; characteristics such as speed, acceleration, flow, density are modelled. At a network level, a road network or a large area is considered; the model is concerned with flows on all the links, or on all the routes (a chain of links connecting two places), or even between each pair of zones (sub-areas of the whole area concerned).

There are two main types of traffic model at the link level, the *car-following model* and the *fluid model.* They both describe the motion of traffic flow on a road stretch, but from different perspectives: the car-following model treats each car individually while the fluid model considers traffic flow as a continuum. For this reason, the car-following model can be regarded as a microscopic model and the fluid model as a macroscopic model. The two models are related in that one form of the fluid model can be derived from the car-following model (Payne, 1979).

At a network level, traffic flows are often referred to as *trips.* Two types of model are relevant to the dynamics of network traffic flow: *trip* (or *traffic*) *distribution models* and *trip* (or *traffic*) *assignment models.* These two types of model are normally used in traffic network planning and management to estimate the flow between each pair of zones of an area, or the traffic flow on each link of a road network. The four models and their relationships (the relationship of the car-following model to the fluid model, and that of the trip distribution models to trip assignment models) are described in detail below.

2.2. THE CAR-FOLLOWING MODEL

2.2.1. The model

The *car-following model* was developed to model the motion of vehicles following each other on a single lane without overtaking (for example, Wilhelm and Schmidt, 1973, Leutzbach, 1988). The model is based on the assumption that a driver responds to the motion of the vehicle immediately in front.

Consider a line of cars numbered from 1 (the leading car) to N (the last car), shown in Figure 2.1. Let $x_n(t)$ denote the position of car n at time t. Then the derivatives $\dot{x}_n(t)$ and $\ddot{x}_n(t)$ are the corresponding speed and acceleration.

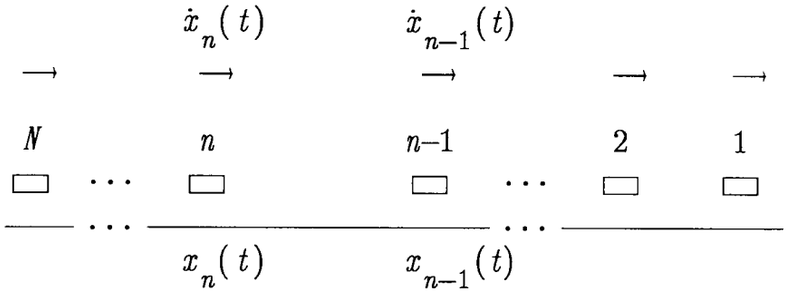


Figure 2.1 The car-following model

It is assumed that the response of a driver to the vehicle in front is to accelerate (or decelerate) his or her vehicle. The resulting acceleration is assumed to equal the *stimulus* multiplied by a *sensitivity*. The stimulus is represented by the difference between the driver's speed and the speed of the vehicle in front and the sensitivity is a measure of the intensity of the response. A larger sensitivity leads to a greater change in acceleration for a given change in stimulus. In addition, a time delay or the *reaction time* is built into the model: the driver does not react immediately to changes in relative speed or spacing. Thus the equations of the car-following model are:

$$\ddot{x}_n(t) = \beta_n \left(\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)\right), \quad n = 2, 3, \ldots, N, \tag{2.1a}$$

where β_n is the sensitivity of the nth car, and τ is a constant, representing the reaction time of the driver of the nth car. The sensitivity depends both on

the current speed of car n and on its distance from car $n-1$ at the time τ previously:

$$\beta_n = \alpha \frac{(\dot{x}_n(t))^m}{(x_{n-1}(t-\tau) - x_n(t-\tau))^l}, \tag{2.1b}$$

where α is a positive parameter; m and l are non-negative parameters, not necessarily integers. When $m = 0$ and $l = 0$, the sensitivity is constant and the model is linear:

$$\ddot{x}_n(t) = \alpha\,(\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)), \quad n = 2, 3, \ldots, N. \tag{2.2}$$

Otherwise, the model is nonlinear.

2.2.2. Literature review

There have been many previous studies of the dynamics of the car-following model. Chandler *et al.* (1958) and Herman *et al.* (1959) examined the linear car-following model (2.2). Two types of stability were considered: local stability and asymptotic stability. *Local stability* is concerned with the time variations of the response of one car to a change in the motion of the car in front. *Asymptotic stability*, on the other hand, is concerned with the manner in which a fluctuation of the motion of the lead car is propagated down a line of traffic. Both forms of stability are defined in terms of safe driving: nonoscillatory, damped response represents the safest driving and is said to be stable, while oscillatory response with or without damping represents hazardous following conditions and is said to be unstable. Conditions for local stability were established using Laplace transforms and numerical analysis of singularities of the inversions of the Laplace transforms (Herman *et al.*, 1959). The qualitative properties of solutions of (2.2) were found to be

(1) if $\alpha\tau > \pi/2$, the solution is oscillatory with increasing amplitude;
(2) if $\alpha\tau = \pi/2$, the solution is oscillatory with constant amplitude;
(3) if $1/\mathrm{e} < \alpha\tau < \pi/2$, the solution is oscillatory with damped amplitude;
(4) if $\alpha\tau \leq 1/\mathrm{e}$, the solution is non-oscillatory and damped.

The asymptotic stability of (2.2) was examined by Fourier analysis and Laplace transforms. The condition for a fluctuation to be damped as it is propagated down a line of cars was found to be $\alpha\tau < 1/2$.

Unwin *et al.* (1967) investigated the so called reciprocal-spacing car-following model, that is, the model of (2.1) with $m = 0$ and $l = 1$:

$$\ddot{x}_n(t) = \alpha \frac{\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)}{x_{n-1}(t-\tau) - x_n(t-\tau)}, \quad n = 2, 3, \ldots, N.$$

This model is nonlinear. Only two cars were considered in their study with the motion of the first car being treated as a forcing term. The delayed terms in the model

$$(\dot{x}_1(t-\tau) - \dot{x}_2(t-\tau)) \quad \text{and} \quad (x_1(t-\tau) - x_2(t-\tau))$$

were expanded by Taylor series around t, so that the model was converted to a second-order ordinary differential equation by taking the first two terms of the Taylor series. The motion of the first car is assumed to be given. A following car was said to be stable with respect to a leading car in the sense of Liapunov if a small change in the initial conditions produces only a small change throughout the solution. The stability was investigated by the Liapunov direct method (Driver, 1977).

There was not much work on the car-following model during the 1970s and 1980s. More recently, there have been two studies about dynamic and possible chaotic behaviour in the car-following model. Disbro and Frame (1990) claimed that chaos can definitely occur in the car-following model. Their conclusion was based on showing that the first Liapunov exponent is positive for certain values of the parameters. However, the existence of a positive Liapunov exponent is a necessary but not a sufficient condition for chaos (The concept of Liapunov exponent will be explained in Chapter 5) and the authors give no plots or other evidence of a strange attractor. On the other hand, Kirby and Smith (1991), in an exploratory study, found no evidence of chaos in the car-following model, though some features of non-linear dynamics related to chaos were found.

Ferrari (1994) considered the stability of a linear two car model

$$\ddot{x}_2(t) = \alpha\,(\dot{x}_1(t-\tau) - \dot{x}_2(t-\tau)).$$

His research is in the same line as that of Chandler *et al.* (1958) and Herman *et al.* (1959) mentioned above. Instead of α and τ, he considered the influence of the characteristics of the speed of the lead car on asymptotic stability.

2.3. THE FLUID MODEL

2.3.1. The model

In contrast to the car-following model, the fluid model describes traffic flow on a road link in an aggregated way by an analogy with the theory of fluids and uses such quantities as flow, density, speed, etc. The fluid model was first suggested by Lighthill and Whitham (1955). Later developments include applications of statistical mechanics and hydrodynamic theory (Payne, 1979, Baker 1983, Leutzbach, 1988). Strictly speaking, models derived from statistical mechanics should be regarded as mesoscopic (the scale between microscopic and macroscopic) models rather than macroscopic ones as they are normally considered to be. The fluid theory is a system of (dynamic) equations describing the interrelationships of traffic flow, density, and speed. The definitions of the three variables are as follows.

(1) *Traffic flow* Q — the number of cars passing a particular point of the road per unit time.

(2) *Density* K — the number of cars per unit length of road at a point in time.

(3) *Speed* V — the speed of the traffic stream.

Normally these quantities all vary with both time and space; in the case of a traffic lane, space is only one dimensional. Thus Q, K, and V are functions of time t and the space coordinate x. The first equation in the fluid model is inherent in the definitions of Q, K, and V: if a traffic flow of density K travels at speed V, then the number of cars passing a fixed point of road per unit time or the flow is given by

$$Q = K\,V. \tag{2.3}$$

The second equation is the continuity equation

$$\frac{\partial K}{\partial t} + \frac{\partial Q}{\partial x} = 0, \tag{2.4}$$

which is derived from the law of conservation of cars in a section of a road. In fluid dynamics, the third equation would be a law of motion derived from Newton's equations. There is no "law of motion" for traffic flow, so some model has to be postulated. Several forms of the third equation have been suggested to complement the above two equations. In the classic fluid model (Lighthill and Whitham, 1955) the third equation is a steady-state relation of speed and density

$$V = U(K).$$

This equation together with (2.3) and (2.4) can describe some traffic phenomena such as kinematic waves in traffic flow. It can not, however, describe instabilities in traffic flow and is not suitable for dynamic analysis. Some dynamic models have therefore been developed by adding an acceleration equation to (2.3) and (2.4) in order to consider the time variations of speed. Different versions of the acceleration equation have been considered (Payne, 1979, Leutzbach, 1988, Ross, 1988, Kühne, 1987, 1991, and Kühne and Beckschulte, 1993). Most of them are of similar form. A typical one is (Payne, 1979):

$$\frac{\mathrm{d}V}{\mathrm{d}t} \equiv \frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = \frac{1}{T}[U(K)-V] - \frac{\nu}{T}K_x\,K, \tag{2.5}$$

where $U(K)$ is a steady-state relation of speed and density, T and ν are constants, and $K_x = \frac{\partial K}{\partial x}$. This model was "derived" by postulating terms which might affect drivers acceleration in different ways. The two terms on the right hand side are called the *relaxation term* and the *anticipation term* respectively. The relaxation term considers the effect of drivers adjusting their speeds to conform to the spacing $1/K$, with the reaction time T (not necessarily the same to the one in the car-following model); and the anticipation term considers the effect of drivers reacting to conditions downstream, for example, to slow down in anticipation of higher densities downstream ($K_x>0$).

2.3.2. Literature review

Most investigations into the fluid model have concentrated on its development, its validation, and methods for solving it numerically. Not very much has been done about the analysis of the dynamics of this model. Kühne (1987) considered the following acceleration equation:

$$\frac{\mathrm{d}V}{\mathrm{d}t} = \frac{\partial V}{\partial t} + V\frac{\partial V}{\partial x} = \frac{1}{T}[U(K)-V] - c_0^2\frac{K_x}{K} + \nu_0\frac{\partial^2 V}{\partial x^2}, \tag{2.6}$$

where c_0 and ν_0 are constants. The first two terms on the right hand side of the equation are also referred to as the relaxation term and the anticipation term respectively by the author. The third term is added by the author and is called a *viscosity term.*

The fluid model using acceleration equation (2.6) together with (2.3) and (2.4) is analyzed by Kühne (1991) and Kühne and Beckschulte (1993). The model was converted to a system of three first-order ordinary differential equations by truncated expansion. The stability of a fixed point and a periodic solution found in the converted equations was analysed numerically. Liapunov exponents were calculated for the periodic solution (Kühne and Beckschulte, 1993) and for some observed speed data (Kühne, 1991). A solution which seems to be irregular was found in the converted model, but the authors did not conclude whether the irregular solution is chaotic or not.

2.4. THE CAR-FOLLOWING MODEL VERSUS THE FLUID MODEL

The car-following model and the fluid model both describe the motion of traffic flow on a road link, though in different ways. In this thesis the car-following model is chosen for the consideration of traffic dynamics at the link level for the following three reasons:

(1) The car-following model is easier to deal with than the fluid model. The fluid model is a system of partial differential equations. Direct analysis of dynamics in partial differential equations is difficult because it inevitably involves the interplay of temporal and spatial behaviour and is normally

very time-consuming in a numerical analysis. Therefore, partial differential equations are often reduced to a system of ordinary differential equations in dynamic considerations. For example, the well-known Lorenz model (Lorenz, 1963) was obtained by truncating the partial differential equation modelling thermal convection between two infinite planes. Kühne (1991) and Kühne and Beckschulte (1993) in their study of the fluid model also reduced the model to a system of ordinary differential equations.

(2) Many stability analyses of traffic flow on a road link have been based on the car-following model. On the contrary, very little progress has been reported of the dynamic behaviour on the fluid model.

(3) Thirdly, there has been some research into the automatic control of driving (See for example, Bender and Fenton,1970, Kapur,1971). In these control systems each vehicle has a built-in driver aid and the vehicles are monitored by an electronic control system. The car-following model can provide a basis for these automatic control systems. Therefore, it is important to understand the dynamics of the model.

The car-following model will be investigated in Chapter 4.

2.5. BRIEF DESCRIPTION OF THE TRAFFIC PLANNING PROCESS

Because the trip distribution model and the trip assignment model can both be used in traffic planning studies, it is worthwhile to give a brief description of the traffic planning process here. The purpose of traffic planning is to determine whether a new stretch of road should be built or whether a current road or road network needs to be improved by, for example, widening or introducing new traffic control strategies. A traffic planning process normally consists of three main steps: (1) a traffic survey, which investigates the current traffic conditions, road conditions, and other relevant socio-economic conditions; (2) traffic forecasting, which predicts the future traffic flows and their distributions on the road network in the study area based on the data from the traffic survey; (3) assessment, which examines the current road network to see if it can meet the needs of the predicted future traffic and, if necessary, to propose a new road network for the future. The steps (2) and (3) are often repeated because if the

road network is changed in step (3) then traffic flow distribution over the network needs to be predicted again.

In traffic planning studies, the whole planning area under consideration is divided into smaller areas called *zones*. A journey from one zone to another is defined as a *trip*. The zone from which a trip starts is called an *origin* and the zone at which a trip terminates is called a *destination*. A particular zone can be both an origin and a destination if trips can both start and terminate there. The traffic survey is also called an *origin-destination survey* or simply an *O−D survey* . In transport modelling a zone is normally represented by a single point called its *centroid*. Trips from and to a zone are assumed concentrated at the centroid.

There are four models used in step (2) mentioned above. These predict, by four successive steps, the distribution of traffic flow on a road network. (a) *Trip generation*, in which the numbers of trips generated from and attracted to each zone are determined given the socio-economic data in the planning area. (b) *Trip distribution*, in which the number of trips between each O–D pair is estimated. (c) *Modal split*, which splits trip makers into the alternative transport modes available, such as between private cars and public buses. (d) *Trip assignment*, which assigns trips between each O–D pair to alternative routes connecting these O–D pairs, so that the traffic flows on each road link can be obtained. Several points need to be addressed about the four steps. First, this may not be the only sequence. Some approaches have put the modal split before the trip distribution. Second, this sequence could also be iterative because some factors such as travel cost are unknown at an earlier stage (trip distribution) and have to be assumed. After the trip assignment is made, the assumed values of costs may need to be modified according to the assigned traffic flow. Third, now there are some approaches which combine two, or three, or even four steps into one step. Fourth, these four models need to be calibrated or validated from traffic survey data to determine the forms of model and values of parameters in the model so that they are suitable to the particular area considered. Fifth, the applications of the trip distribution model and trip assignment model are not limited to the traffic planning process. They are also used in traffic management schemes to estimate or to predict flows on a road network.

The trip distribution model and the trip assignment model will be considered in this thesis. Although it might be possible to examine dynamic versions of trip generation models and modal split models, these two types of model will not be

considered here. In modal split models typically there are only two choice alternatives, that is, public transport and private transport. Therefore the model does not appear to be promising for interesting dynamic properties. As for trip generation models, the research may be in a much more general context, including socio-economic developments and land uses, than that of the dynamic behaviour of traffic flows. In addition, there is not enough time to include these two types of models in the thesis. The trip distribution model and the trip assignment model will be reviewed in the next two sections respectively.

2.6. THE TRIP DISTRIBUTION MODEL

2.6.1. Definitions and notation

It is now customary to represent the trip pattern in a study area by means of a *trip matrix*. It is essentially a two-dimensional array of cells where rows and columns represent each of the zones in the area, as shown in Table 2.1.

Table 2.1 A general form of a trip matrix

		destination						total
		1	2	....	j	...	J	$\Sigma_j T_{ij}$
origin								
1		T_{11}	T_{12}	...	T_{1j}	...	T_{1J}	O_1
2		T_{21}	T_{22}	...	T_{2j}	...	T_{2J}	O_2
$\vdots$								
i		T_{i1}	T_{i2}	...	T_{ij}	...	T_{iJ}	O_i
$\vdots$								
I		T_{I1}	T_{I2}	...	T_{Ij}	...	T_{IJ}	O_I
total	$\Sigma_i T_{ij}$	D_1	D_2	...	D_j	...	D_J	$\Sigma_{ij} T_{ij} = T$

In this table, T_{ij} is the number of trips from zone i to zone j, O_i is the total number of trips originating in zone i, D_j is the total number of trips attracted to zone j, T is the total number of trips from all origins or to all destinations, I

is the number of origins, and J is the number of destinations. The sum of trips in the ith row should equal O_i. The sum of trips in jth column should equal D_j. And the sum of all elements in the matrix should equal T. These marginal constraints can be written as:

$$\Sigma_j T_{ij} = O_i, \qquad i = 1, 2, ..., I, \tag{2.7a}$$
$$\Sigma_i T_{ij} = D_j, \qquad j = 1, 2, ..., J, \tag{2.7b}$$
$$\Sigma_{ij} T_{ij} = T. \tag{2.7c}$$

These constraints are expressed in terms of absolute values. In trip distribution modelling, trips are often normalized so that relative quantities can be used. Let

$$t_{ij} = T_{ij}/T,$$
$$o_i = O_i/T,$$
$$d_j = D_j/T,$$

where o_i is the total relative number of trips originating from zone i, and d_j is the total relative number of trips terminating at zone j. Then the constraints in (2.7) become

$$\Sigma_j t_{ij} = o_i, \qquad i = 1, 2, ..., I, \tag{2.8a}$$
$$\Sigma_i t_{ij} = d_j, \qquad j = 1, 2, ..., J, \tag{2.8b}$$
$$\Sigma_{ij} t_{ij} = 1. \tag{2.8c}$$

The O_i's and D_j's are often obtained from trip generation models and are used as inputs to trip distribution models, which in turn give the T_{ij}'s, the entries of a trip matrix. If a trip distribution model satisfies both (2.7a) and (2.7b), in other words, the total number of trips originating and terminating in each zone given by the model equals the predetermined O_i's and D_j's respectively, then the model is said to be *doubly constrained.* If a model satisfies (2.7a) or (2.7b) but not both then it is *origin constrained* or *destination constrained.* The origin constrained and destination constrained models are also called *singly constrained* models. If a model satisfies (2.7c) only it is called *unconstrained* model. What kind of constraints should be involved in a model depends mainly on the availability of information about the O_i's and D_j's. Other considerations can also be given in determining the kind of constraint to be included. For example, for work trips, both trip origins (for example, residence locations) and destinations (for example, work places) are based on long-term decisions.

Consequently, the total flow from and to each zone tends to be fixed and a doubly constrained model is appropriate. For non-work trips, the total number of trips from an origin may be regarded as fixed while the total number of trips to each destination is a result of daily travel choice of trip makers and should be considered to be variable. In this case an origin-constrained model is more suitable. That the total flow attracted to each zone is known does not necessarily mean that a doubly constrained model should be used.

Associated with each O–D pair there is a travel cost. It may be considered in terms of distance, time or monetary units, or a combination of these. It represents the disutility of a journey and is normally referred to as the *generalized cost of travel.* Normally the travel cost between each O–D pair increases with the traffic flow between the O–D pair. Travel cost is an important factor in determining the distributions of trips in a trip matrix.

In trip distribution models the number of trips between each pair of zones is estimated on the basis of any information available, such as the attractivity of destinations, the productivity of origins, and travel costs between O–D pairs. Different trip distribution models have been developed for different sets of problems and conditions. There are two types of model. The first type of model is the *growth factor model.* It is used for updating a trip matrix or for forecasting a future trip matrix. The information available here includes a basic trip matrix, perhaps obtained from a previous study or from a recent O–D survey, and some growth factors. It is assumed that the future number of trips between each pair of zones is the current number of trips multiplied by a growth factor. This type of model is irrelevant here. The second type of model is the *synthetic model.* These models estimate (often for the future) the number of trips between each pair of zones without directly using a known trip pattern. Instead, they consider trip-making behaviour and the way this is influenced by such factors as total number of trips in each zone and costs involved. Therefore, they are called synthetic models as opposed to growth factor models. The best known of these models is the *gravity model.* This model is of interest here and is described in the next subsection.

2.6.2. The gravity model

The aim of gravity models is to estimate the number of trips between each O–D

pair based on travel costs between zones and the total number of trips from and/or to each zone, or the total number of trips in the whole area under consideration. A family of gravity models in terms of relative quantities can be written in a general form

$$t_{ij} = \psi\, f(c_{ij}), \quad i = 1, 2, \ldots, I, \;\; j = 1, 2, \ldots, J, \tag{2.9}$$

where t_{ij} is the relative number of trips from zone i to zone j, normalized so that $\Sigma_{ij} t_{ij} = 1$, c_{ij} is the corresponding travel cost, $f(c_{ij})$ is called the *deterrence function* which relates the number of trips to the travel costs, and ψ is an appropriate normalizing factor which may depend on i and/or j.

Three types of deterrence function are used in practice (Ortúzar and Willumsen, 1990): (a) exponential function, (b) power function, and (c) combined function. They can be written as:

$$f(c_{ij}) = c_{ij}^{\mu} \exp(-\beta\, c_{ij}),$$

where μ and β are constants. When $\mu=0$ and $\beta>0$, f is an exponential function; when $\mu<0$ and $\beta=0$ it is a power function; and when $\mu>0$ and $\beta>0$ it is a combined function. In the first two forms the number of trips is a decreasing function of cost, while in the third the number of trips first increases and then declines as cost increases; the location of the turning point of the function depends on the relative magnitude of μ and β.

The factor ψ in (2.9) is chosen so that the appropriate constraints (2.8a)–(2.8c) of an O–D matrix are satisfied. The form of ψ for the unconstrained, singly constrained, and doubly constrained models are as follows.

(1) Unconstrained model. Only (2.8c) is satisfied and

$$\psi = \frac{1}{\Sigma_{ij} f(c_{ij})}, \quad \text{so that} \quad t_{ij} = \frac{f(c_{ij})}{\Sigma_{kl} f(c_{kl})}.$$

(2) Singly constrained model. For an origin-constrained model (2.8a) is met and ψ is replaced by a set of constants:

$$a_i = o_i \frac{1}{\Sigma_j f(c_{ij})}, \quad \text{so that} \quad t_{ij} = o_i \frac{f(c_{ij})}{\Sigma_l f(c_{il})}.$$

For a destination-constrained model (2.8b) is met and ψ is replaced by another set of constants:

$$b_j = d_j \frac{1}{\Sigma_i f(c_{ij})}, \quad \text{so that} \quad t_{ij} = d_j \frac{f(c_{ij})}{\Sigma_k f(c_{kj})}.$$

(3) Doubly constrained model. Both (2.8a) and (2.8b) are satisfied and ψ is replaced by two sets of constants, or the *balancing factors*:

$$a_i = o_i \frac{1}{\Sigma_j b_j f(c_{ij})}, \quad b_j = d_j \frac{1}{\Sigma_i a_i f(c_{ij})}.$$

And

$$t_{ij} = a_i b_j f(c_{ij}).$$

In the doubly constrained model, the two sets of balancing factors depend on each other; calculations of t_{ij} need a special iterative method. This will be described in Chapter 5.

2.6.3. A dynamic gravity model

Most gravity models in the literature are static and the travel costs are assumed to be independent of the number of trips. A dynamic model is needed to study the O–D flow dynamics in an area. Dendrinos and Sonis (1990) proposed an iterative version of the gravity model by assuming that the future number of trips depends on the current travel cost which, in turn, is a function of the current number of trips. That is,

$$t_{ij}(n+1) = \psi(n)\, f(c_{ij}(n)), \tag{2.10a}$$

$$c_{ij}(n) = c^0_{ij} \left[1 + \alpha \left[\frac{t_{ij}(n)}{q_{ij}} \right]^{\gamma} \right], \tag{2.10b}$$

where c^0_{ij} is the uncongested travel cost from zone i to zone j, q_{ij} is the

corresponding capacity of the roads (the ability that roads can accommodate traffic flows), and α, and γ are positive constants. For example, the dynamic unconstrained gravity model can be written as

$$t_{ij}(n+1) = \frac{f(c_{ij}(n))}{\Sigma_{kl} f(c_{kl}(n))}$$

with $c_{ij}(n)$ being given by (2.10b).

Dendrinos and Sonis suggested that chaotic behaviour may exist in this model, but did not give any further results. The dynamic model (2.10) will be investigated in Chapter 5. Although they suggested the particular form of function for $c_{ij}(n)$ in (2.10b), other increasing functions may also be used. This will be considered in Chapter 5.

2.7. THE TRIP ASSIGNMENT MODEL

In the trip distribution process, the number of trips between each pair of zones is determined without specifying on which particular route these trips are made. In the modal split process, trips are allocated to different travel modes, normally between private transport mode and public transport mode. It is the task of trip assignment models to allocate trips between zones to one or more particular routes connecting the zones. A route can be a chain of road links of a road network in the case of private transport, or a route of a public transport service in the case of public transport. Normally, trip assignment is carried out for private transport and public transport separately; in the former traffic flows are assigned to a road network and in the latter passengers are assigned to a transit network. In this project, only private transport assignment is considered. Therefore, the trip assignment here means to load a trip matrix of road traffic onto a road network.

Trip assignment models are more complicated than trip distribution models because they deal with the spatial distributions of traffic flows on a more detailed scale. The representation of a road network is introduced in the next subsection. Trip assignment models are described and reviewed in the subsequent two subsections. Finally, a dynamic model to be investigated in this thesis is outlined in the last subsection.

2.7.1. Representation of a road network

A typical road network consists of junctions and road links through which traffic moves. It can be represented naturally by a *directed graph*, that is, a system of *nodes* and *links* joining them. A node represents a junction in most cases and a link a homogeneous stretch of road between junctions. A node can also be a centroid, representing an origin or a destination. These special nodes can be considered as "source" and "sink" nodes where trips originate and terminate. Each centroid is attached to the road network by one or more *dummy links* which represent the average travel distance of joining the transport system for trips from and to that zone. These dummy links are chosen to connect a centroid to one or more neighboring nodes of the network. A *route* is a chain of links from an origin to a destination. An O–D pair may be connected by more than one route. In fact, in a large road network, the number of routes connecting an O–D pair can be so large that it is often hardly possible to enumerate all of them explicitly. Many trip assignment algorithms and software packages are therefore link based procedures, though others are route based.

In a road network, each link has a travel cost associated with it: *link cost.* This can represent travel time, distance, etc. Normally a *generalized cost* is used. A link cost depends on the link length, link capacity (the ability of a link to accommodate traffic flow), and above all, traffic flow. In some traffic assignment models congestion is not considered and link costs are assumed to be independent of link flows. When the congestion effect is considered, the link cost is assumed to be a function of the link flow, called a *link performance function.* It is normally assumed to be a monotonically increasing function. Several types of function have been used (Branston, 1976). Clearly, the travel cost on each route is the sum of the costs on all links comprising that route; the flow on each link is the sum of flows on all routes using that link.

2.7.2. Trip assignment models

Given a trip matrix, and the layout and characteristics such as travel costs of a road network, trip assignment models allocate trips to alternative routes. The outcome of a trip assignment consists of flows on each route and each link, or a

flow pattern in the road network. The main factor affecting the assignment is drivers' *route choice* behaviour. Although drivers tend to choose the cheapest route, different drivers travelling between the same two points often choose different routes. This fact can be accounted for by two types of reason. (1) Stochastic effects in route choice: different drivers may have different definitions and perceptions of the "best" or the "cheapest" route and hence may choose different routes; (2) capacity constraints: congestion makes an initially cheaper route less attractive and causes drivers to switch to alternative routes.

Trip assignment models can be roughly categorized as shown in Table 2.2 according to whether or not the stochastic effect and the congestion effect are considered.

Table 2.2 Classification of trip assignment models

		Stochastic Effects Included?	
		No	Yes
Capacity Constraint Included?	No	All-or-nothing assignment	Multi-routeing (or stochastic) assignment
	Yes	User equilibrium assignment	Stochastic user equilibrium assignment

Trip assignment models which do not consider stochastic factors in route choice are deterministic. Deterministic models assume that drivers are homogeneous and have perfect knowledge of road and traffic conditions, and hence travel costs. Route choice in these models is based on real or *measured* travel costs. Stochastic assignment methods, on the other hand, allow for variations in drivers' judgments of travel costs, and route choice is based on *perceived* travel costs. The perceived travel costs are random variables. Different stochastic assignment models differ in the assumption about the distribution of the random component in the perceived cost. There are two main types of stochastic assignment method: the logit-based method and the probit-based method. The random term of the perceived cost is assumed to be an independently and identically distributed Gumbel variable in the logit model and a normally distributed variable in the probit model. The logit model is much easier to use because the

choice probability (the probability that a particular route will be chosen) can be expressed analytically. With the probit model, however, the choice probability can only be approximated analytically. Alternatively, it can be obtained by computer simulations. These methods are very time consuming, especially for a general network in which the numbers of routes between O–D pairs can be very large.

If the congestion effect is not considered then trips are assigned to a road network according to a set of predetermined link travel costs. The all-or-nothing method assigns all trips between each O–D pair to the cheapest or the shortest route; while the multi-routeing method splits flows between alternative routes according to perceived costs. In both cases, some links may be overloaded and may become congested. Equilibrium assignment methods take congestion effects into account by assuming that link costs depend on link flows. Thus congested links will become more expensive to travel and flows may be diverted to other routes. A user equilibrium (UE) or a stochastic user equilibrium (SUE) is reached when the cost for every traveller is the minimum (in UE assignment) or is thought to be the minimum (in SUE assignment). The UE and the SUE will be explained below. It should be pointed out that the route or link flows used to define the SUE are deterministic variables. Therefore, the SUE assignment models are deterministic in nature although it is often referred to as "stochastic" models. The choice probability in these models should be considered as the choice proportion, that is, the fraction of the trip between an O–D pair that uses a particular route. The equilibrium assignment models are of interest here and a literature review is given below.

2.7.3. Literature review

There have been many studies of equilibrium assignment in the literature (See, for example, Fernandez and Friesz, 1983, and Friesz, 1985 for reviews). Almost all these studies have concentrated on finding an equilibrium solution first proposed by Wardrop (1952). The equilibrium is a flow pattern on a road network that satisfies conditions derived from certain behavioural principles of drivers. One of the Wardrop equilibria is the *user equilibrium.* It is a flow pattern under which no driver can reduce his or her travel cost by changing a route. Earlier methods for finding the user equilibrium were heuristic. The two most commonly used heuristic methods are the *capacity restraint* method and the

incremental assignment technique (Sheffi, 1985). Both of them are iterative methods. It is often found, however, that the simple heuristic iterative methods may not converge because of oscillations, or even when they do converge, they may produce a flow pattern which is different from the user equilibrium.

More recently, the problem of finding the user equilibrium has been formulated as other equivalent problems, including nonlinear complementary problems, nonlinear variational inequality problems, and mathematical programming problems (Friesz, 1985). In these approaches the user equilibrium is defined and formulated. The existence and the uniqueness of the equilibrium is established. Then an algorithm is developed to solve the problem. Among these methods, the most widely used method is to formulate the equilibrium assignment problem as a mathematical programming problem whose solution gives the user equilibrium. All these methods are static: they deal with only a steady-state of traffic flows. The time variation of flow patterns in the network is not involved.

It was not until the 1980s that attempts were made towards understanding dynamic network equilibrium. There are two approaches to the problem of dynamic trip assignment. One is to consider explicitly the time-dependence of road network characteristics, such as travel costs, and traffic flows. Dynamic link performance functions are used to consider queueing and congestion effects on links; the distribution of departure times of O–D trips is assumed to depend on the temporal distribution of travel costs over the network so that the O–D flows vary with time. Another type of dynamic approach is to model the process of adjustment of the flow pattern in a network from one time instant to another for a given O–D matrix. This second type of study has received much less attention and is in the same line of research as that of this thesis.

Horowitz (1984) investigated the stability of a "stochastic" equilibrium in a discrete-time assignment model for a network of one O–D pair connected by two links. It is assumed that in each time period, drivers choose the link that is currently *perceived* to be cheaper. The perceived cost in each period of time is taken to be a weighted average of the costs in previous time periods. Three different ways of weighting are proposed. Horowitz proves that the dynamic process of adjustment will approach the stochastic user equilibrium if the relative weights for the costs of recent past and the distant past are properly balanced. A deterministic model is also mentioned, which is the same as the "stochastic" model except that drivers know perfectly the link costs and choose the link that

is *measured* to be cheaper. He indicates that the conditions of stability for the "stochastic" model do not suit the deterministic model and so the stability of the equilibrium can not be assured for the deterministic model.

Smith (1984) proposed a continuous-time adjustment mechanism modelled by a system of ordinary differential equations. He assumed that the rate of change of route flows depends on route flows and on relative costs on alternative routes. The equilibrium of this dynamical system coincides with the user equilibrium. Using a method due to Liapunov, Smith was able to prove that the equilibrium is stable if the cost-flow function (functions describing the dependence of route costs on route flows) is monotonic and smooth.

It can be seen that all the studies, static, or dynamic, seek to obtain a solution representing the user equilibrium. It is somewhat amazing that, as Friesz (1985) pointed out, "despite the fact user equilibrium has been employed as the key behavioral assumption in most urban transportation network models, little effort has been expended to determine whether real world network flow patterns are actually described as user equilibria".

2.7.4. A dynamic logit-based trip assignment model

To consider the dynamics of trip assignment, a dynamic assignment model is essential. The kind of dynamic model suitable here is that which describes, for a given trip matrix and a road network, how the flow pattern in the network changes with time. Therefore, the second type of dynamic consideration (to consider the time adjustment of one flow pattern to another) is relevant here. Unfortunately, very little experience or knowledge is available to determine even a plausible adjustment mechanism. Here we base the consideration of the dynamics in trip assignment on the logit assignment model. The model is modified for dynamic considerations. The modified model considers both drivers' differences in route choice and congestion effects, as we shall see below.

The logit-based model was first suggested by Dial (1971). It is assumed that an individual driver chooses alternative routes according to the route cost in the way modelled by the logit discrete choice model. For a network of one O–D pair connected by two links, for example, the model can be written as

$$x_i = \frac{\exp(-\theta c_i)}{\Sigma_j \exp(-\theta c_j)}, \quad i = 1, 2,$$

where x_i is the assigned flow on route i, c_i is the travel cost on route i, and θ is a positive constant. The flows in this equation have been normalized so that $x_1 + x_2 = 1$. This model is static and does not consider congestion effects. An iterative dynamic model can be constructed in the similar way to that for the gravity model. The cost on each link can be associated with the flow on that link by the link performance function; the link flows can be calculated from the link costs in the previous time period. That is

$$x_i(n+1) = \frac{\exp(-\theta c_i(n))}{\Sigma_j \exp(-\theta c_j(n))}, \quad i = 1, 2, \tag{2.11a}$$

and

$$c_i(n) = g_i(x_i(n)), \quad i = 1, 2, \tag{2.11b}$$

where $x_i(n)$ is the flow on route i at the nth time period, $c_i(n)$ is the travel cost on route i at the nth iteration, $g_i(x_i)$ is the link performance function, and θ is a positive constant. Dendrinos and Sonis (1990) actually suggested the idea of this kind of extension, though to a general individual discrete-choice model and without any further analysis. The dynamic model for a general road network is much more complicated than the one for a two-link network. The route choices are route based while the link performance function is link based. These will be discussed in more detail in Chapter 6, where the dynamic assignment model will be examined.

2.8. SUMMARY

It has been proposed that the following three traffic models will be examined to study the dynamics of traffic flow.

(1) The car-following model (2.1), which is a system of delay-differential equations modelling a line of cars moving on a stretch of road.

(2) The dynamic gravity model (2.10a) and (2.10b) suggested by Dendrinos and Sonis (1990), which is a iterative model describing the dynamics in trip distribution process.

(3) The logit-based dynamic trip assignment model (2.11a) and (2.11b) suggested indirectly by Dendrinos and Sonis (1990), which is also a iterative model describing the dynamics in trip assignment process.

CHAPTER 3. NONLINEAR DYNAMICS: CONCEPTS AND METHODS

This thesis is concerned with the dynamics of traffic models. The variations of traffic flows are investigated in the context of *nonlinear dynamics*, that is, the mathematics of dynamical systems governed by nonlinear equations. In this chapter, some basic concepts and methods of nonlinear dynamics are introduced; methods to be used in the thesis are outlined.

3.1. BASIC CONCEPTS OF NONLINEAR DYNAMICS

3.1.1. Dynamical systems and attractors

A *dynamical system* is a system whose state varies with time. The state of the system is characterized by a system of state variables, $x_1, x_2, \ldots, x_m$, where m is the number of variables. This set of variables is chosen so that it can characterize the state of the system completely and throughout its evolution. One history of the evolution of the system is defined by the time series $\mathbf{x}(t) = [x_1(t), x_2(t), \ldots, x_m(t)]$, $0 \le t < \infty$, the path of which is a *trajectory* in an m dimensional space, or the *phase space* of the system. The trajectory is also called an *orbit*. The time evolutions of a dynamical system can be modelled by a set of evolutionary equations in terms of the state variables. A typical form of evolutionary equations is a system of ordinary differential equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}(t) = \mathbf{F}(\mathbf{x}(t)), \qquad \mathbf{x}(t) \in S, \tag{3.1}$$

where $S \subseteq \mathbb{R}^m$ is the phase space of the system. Given an initial condition $\mathbf{x}_0 \in S$, then a function

$$\mathbf{x}(t) = \varphi_t(\mathbf{x}_0)$$

is a solution of the equation if it satisfies the equation such that

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\varphi_t(\mathbf{x}_0)\right) = \mathbf{F}(\varphi_t(\mathbf{x}_0)),$$

$$\varphi_0(\mathbf{x}_0) = \mathbf{x}_0.$$

The solution is a continuous curve through $\mathbf{x}_0$, which traces out a trajectory in the phase space.

Dynamical systems whose states and evolutions can be described by finite numbers of variables are finite dimensional. The system (3.1) is of this kind. More complicated systems can be modelled by partial differential equations, such as the fluid traffic flow model (2.4), or by delay-differential equations, such as the car-following equation (2.1). These dynamical systems are *infinite dimensional* because the evolutions of the system depend on *initial functions* which involve infinite numbers of initial values.

Dynamical systems defined by differential equations are *continuous systems*: the state of the system varies with time continuously. A different type of dynamical system is the *discrete system*, modelled typically by finite difference equations or iterated maps (functions)

$$\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n)), \qquad \mathbf{x}(n) \in S, \tag{3.2}$$

where n is the discrete time and $S \subseteq \mathbb{R}^m$ is the phase space. In this kind of system, the state of the system changes with time in discrete steps. By iterating the above equation given an initial point $\mathbf{x}(0) \in S$, one gets a solution

$$\mathbf{x}(0), \mathbf{x}(1), \ldots, \mathbf{x}(n), \ldots,$$

which is a sequence of points in the phase space. This sequence of points defines the trajectory through $\mathbf{x}(0)$. The dynamic gravity model (2.10a–2.10b), and the dynamic logit based trip assignment model (2.11a–2.11b) are both discrete systems.

A dynamical system is linear if the evolutionary equations are linear (The equations are linear means that the function $\mathbf{F}$ or $\mathbf{f}$ is a linear function). Otherwise, it is nonlinear. Most dynamical systems in practice are nonlinear. The three traffic models to be studied in this thesis are all nonlinear. Generally

speaking, nonlinear differential or difference equations such as those mentioned above cannot be solved analytically; theoretical analysis of the dynamics can only be made to a limited extent. Many theoretical analyses are made by *ad hoc* methods. On the other hand, these equations can be integrated or iterated numerically to get the solutions or the trajectories, which can then be examined geometrically by time series plots (the solutions against time) or phase portrait plots (the trajectories in the phase space). In fact, these geometrical methods are major tools in the analysis of nonlinear dynamical systems.

Usually, for an arbitrarily chosen initial point $\mathbf{x}(0) \in S$, trajectories of many dynamical systems exhibit first a transient, and then settle down to some form of steady-state solution in a *bounded* subset of the phase space. This subset is called an *attractor* for it "attracts" neighbouring trajectories. If the system starts outside but close enough to the attractor, the trajectory will always converge to the attractor. What is important is the steady-state solutions rather than transients. This is because the steady-state solutions represent the long term behaviour of the system, or the behaviour as time goes to infinity and when initial transients have died away. Steady-state solutions must be bounded in order to make sense. More will be said about steady-state solutions in 3.1.3.

There are basically four types of attractor. The simplest form of an attractor is a *point attractor*. It is a stationary point towards which trajectories go and where motion stops. It is also called a *stable equilibrium* or a *stable fixed point*. A stationary point or an equilibrium $\mathbf{x}^e$ is defined by

$$\mathbf{F}(\mathbf{x}^e) = 0 \qquad \text{in the continuous system (3.1),} \tag{3.3}$$

or by

$$\mathbf{x}^e = \mathbf{f}(\mathbf{x}^e) \qquad \text{in the discrete system (3.2).} \tag{3.4}$$

An equilibrium is not necessarily a point attractor because it may not be stable (See 3.1.3).

The next type of attractor is a *periodic attractor*, on which the system repeats itself perpetually once it reaches the attractor. A periodic attractor is defined by a periodic solution of the evolutionary equations. Such a periodic orbit is topologically a circle in continuous systems, or a finite set of points in discrete systems. The simplest periodic orbit in a discrete system is composed of two points and is called a *period-2 orbit*; the state of the system oscillates between the

two points. More general periodic orbits are a *period-K orbit* containing K points, on which the system visits each point in turn before returning to the first. As for equilibria, a periodic orbit need not be an attractor.

The third type of attractor is a *quasi-periodic attractor*, which is defined by a quasi-periodic function. A *quasi-periodic function* is a function that can be expressed as the sum of a countable number of periodic functions, each of whose frequency is an integer combination of a finite set of linearly independent base frequencies. Topologically, a quasi-periodic attractor is a p-torus, where p is the number of the base frequencies. Unlike a periodic solution, the trajectory of a quasi-periodic solution does not repeat itself exactly. It moves on the surface of the torus and will eventually cover the torus. Again, a quasi-periodic orbit may not be an attractor.

As an example, we shall show a quasi-periodic solution of the forced Van der Pol equation (Parker and Chua, 1989)

$$\dot{x} = y,$$
$$\dot{y} = (1 - x^2)\, y - x + 0.5 \cos (1.1\ t).$$

This system is a little different from the system (3.1) in that the time t occurs explicitly on the right hand side of the equation. This kind of system is said to be *non-autonomous* because the evolution of the system depends on time. While a system like (3.1) whose evolution does not depend on time is *autonomous*. A non-autonomous system can be converted into an autonomous system by introducing an artificial new variable $z = t$. For example, the Van der Pol equation can be written as

$$\dot{x} = y,$$
$$\dot{y} = (1 - x^2)\, y - x + 0.5 \cos (1.1\ z),$$
$$\dot{z} = 1,$$

which is an autonomous system in the three-dimensional space of x, y, and z. However, the solution of this system is unbounded since z will tend to infinity eventually. There cannot be an attractor in this kind of system for an attractor must be bounded. To overcome this, observe that the Van der Pol equation is periodic with period $T = 2\pi/1.1$. In other words, the equations remain the

same if z is replaced by $z+T$. Thus instead of $z = t$, we introduce a periodic function

$$\theta = 2\pi t/T \qquad \text{mod } 2\pi$$

or

$$\theta = 1.1t \qquad \text{mod } 2\pi$$

to the system. In this, X mod 2π means 2π times the fractional part of $X/(2\pi)$. That is

$$X \text{ mod } 2\pi \;=\; X - 2\pi \times [X/(2\pi)],$$

where $[Y]$ denotes the integer part of Y or the greatest integer less than or equal to Y. Then $0 \leq \theta < 2\pi$, and the Van der Pol equation becomes

$$\dot{x} = y,$$
$$\dot{y} = (1 - x^2)\, y - x + 0.5 \cos \theta,$$
$$\dot{\theta} = 1.1.$$

This system is autonomous and is in the cylindrical phase space of x, y, and θ, where θ is measured in radians. Figure 3.1 shows the phase portrait projection in the (x,y)-plane of a quasi-periodic solution of the Van der Pol equation. It can be seen that the trajectory moves within an annular-like region of the phase space.

The final type of attractor is a chaotic attractor which is a new type of attractor. Although the discovery of chaos may be traced back to the beginning of the century due to the research of Henri Poincaré in geometric dynamics, it was not until the past twenty or thirty years that chaos has grown into a major and exciting topic of research. Chaotic attractors and chaos deserve a separate subsection and are described below.

3.1.2. Chaotic attractors

There is not yet a widely accepted definition for chaotic attractors; one may say that a chaotic attractor is neither a point attractor, nor a periodic attractor, nor

a quasi-periodic attractor. However, it suffices, at least for the purpose of this thesis, to state that *chaotic attractors* are attractors on which solutions have *sensitive dependence on initial conditions* (Eckmann and Ruelle, 1985). The sensitive dependence on initial conditions is referred to as *chaos* or *strangeness*. A chaotic attractor is also called a *strange attractor*. A dynamical system is said to be chaotic or to have chaos if it possesses one or more chaotic attractors. Not all dynamical systems modelled by differential or difference equations have chaotic solutions. In particular, chaos cannot occur in linear systems. It may or may not occur in a nonlinear system.

Sensitive dependence on initial conditions means that two initially close trajectories on a chaotic attractor can diverge so quickly that they soon appear very different. Suppose the two trajectories start at slightly different points separated by $\delta\mathbf{x}(0)$, which will become $\delta\mathbf{x}(t)$ at time t. For deterministic systems modelled, for example, by differential equations, one might expect that $\delta\mathbf{x}(t)$ remains small if $\delta\mathbf{x}(0)$ is small. For chaotic attractors, however, the initial separation may grow exponentially:

$$\delta\mathbf{x}(t) \sim \delta\mathbf{x}(0)\mathrm{e}^{\lambda t}, \tag{3.5}$$

where λ is a measure of average rate of change. Of course, if the two trajectories start at precisely the same point, then they will always stay together since the system is *deterministic*. In practice, however, the initial state of a system can only be specified with limited precision. The exponential growth of small separations makes it impossible to predict even the near future of the system; the behaviour appears *stochastic*. Therefore chaos may also be referred to as *apparently stochastic behaviour in deterministic systems*. The time variation of the state of the system is irregular on chaotic attractors; it is regular on non-chaotic attractors.

A chaotic attractor is, nonetheless, an attractor. It attracts trajectories surrounding the attractor. Therefore the exponential divergence of adjacent trajectories on the attractor cannot go on forever: the trajectories will fold back at some point since the attractor is bounded. This exponential divergence and folding back (*stretching and folding*) makes chaotic attractors very complicated geometrically. Chaotic attractors are not simple geometrical objects like a circle or a torus. The trajectories move on the attractor irregularly and without repeating themselves so that they become densely distributed on the complicated

object as time goes by. However, chaotic attractors are not completely unstructured; they have their own patterns.

These features of chaotic attractors can be seen in the following two examples also discussed, for instance, by Thompson and Stewart (1988). Probably the most well-known chaotic attractor is the attractor found in the Lorenz equations (Lorenz, 1963)

$$\dot{x} = -\sigma\,(x-y),$$
$$\dot{y} = -xz + \rho x - y,$$
$$\dot{z} = xy - \beta z,$$

where σ, ρ, and β are positive parameters. This is a three dimensional continuous system. With $\sigma = 10$, $\rho = 28$, and $\beta = 8/3$, the system has a chaotic attractor. Figures 3.2a–3.2b show the time series and the phase portrait projection on the (x,z)-plane of the attractor. To see the sensitivity to initial conditions, two time series with slightly different initial conditions are plotted in Figure 3.2c, where the solid line and the dashed line are the solutions starting at

$$(x,\ y,\ z) = (8.8070,\ 0.7247,\ 35.3542)$$

and

$$(x,\ y,\ z) = (8.8071,\ 0.7247,\ 35.3542)$$

respectively. Both initial conditions are on the attractor, so there is no transient. It can be seen that the solutions seem to stay together for a while and then become completely different.

Another well-known chaotic attractor is that of the Hénon map (Hénon, 1976)

$$x(n+1) = -a\,[x(n)]^2 + y(n) + 1,$$
$$y(n+1) = b\,x(n),$$

where a and b are positive parameters. This is a two dimensional discrete system. There is a chaotic attractor when $a = 1.4$ and $b = 0.3$. Figure 3.3a shows the phase portrait of the chaotic attractor. The sensitive dependence on initial conditions is shown in Figure 3.3b, where the solid line is the solution

starting at

$$(x, y) = (0.6715, 0.1806)$$

and the dashed line starting at

$$(x, y) = (0.6714, 0.1806).$$

Here again, an initial run is made to let the transient pass by and the orbit plotted is on the attractor. The trajectories begin to separate after only 10 iterations.

The sensitive dependence on initial conditions can be characterized by *Liapunov exponents*, which measure the average rate of growth of small separations on an attractor. An intuitive definition of a Liapunov exponent may be given by λ in (3.5). The definition will be refined. A positive Liapunov exponent indicates expansion of nearby trajectories and a negative exponent contraction of nearby trajectories. Chaotic attractors are characterized by positive Liapunov exponent(s). Non-chaotic attractors do not have positive Liapunov exponents. Typically (although not necessarily), chaotic attractors are *fractals*, or complicated geometric objects which can be regarded as having a fractional (non-integer) dimension (Parker and Chua, 1989). Liapunov exponents and fractal dimensions will be described in detail in Chapter 5, where both measures will be calculated for chaotic attractors found in the gravity model.

3.1.3. Stability

We have referred to the long-term behaviour of a dynamical system as the steady state, though the motion of the system need not be time-independent, except in the case of an equilibrium state. In general, a steady-state solution is *recurrent*: that is, it will return arbitrarily close to its initial point an infinite number of times in the future. Such a recurrent solution may be an attractor. However, not all recurrent solutions attract; some of them may repel in at least one direction: the orbit of a point close to the solution may move away from it. In other words, the solution may not be stable. Stability of a steady-state solution is concerned with the robustness of the solution to small disturbances. Roughly speaking, a steady-state solution is stable if all trajectories starting nearby stay

nearby. Otherwise, if at least one nearby trajectory moves away from the steady-state solution, then the steady-state solution is unstable. The definition can be formalized.

An equilibrium $\mathbf{x}^e$ is *stable* if, for every open neighborhood Ω of $\mathbf{x}^e$ in the phase space, there is a smaller neighbourhood Ω_1 of $\mathbf{x}^e$ in Ω, such that every trajectory that starts in Ω_1 remains in Ω for all future time. In addition, the equilibrium is *asymptotically stable* if there is an open neighbourhood Ω_0 of $\mathbf{x}^e$ in the phase space such that every trajectory starting in Ω_0 tends to $\mathbf{x}^e$ as time tends to infinity. An equilibrium which is not stable is said to be *unstable*. This means that there is at least one trajectory starting arbitrarily close to $\mathbf{x}^e$ which goes away from $\mathbf{x}^e$. The largest possible Ω_0 for which all solutions tend to $\mathbf{x}^e$ is called the *basin of attraction* of $\mathbf{x}^e$; any trajectory starting from a point in the basin of attraction converges to the equilibrium. If the basin of attraction of $\mathbf{x}^e$ is the entire phase space, then the equilibrium is *globally asymptotically stable*.

In practice, however, the stability or the asymptotic stability of an equilibrium in a nonlinear dynamical system may be investigated only *locally* by linearization. Direct analysis of the stability of an equilibrium in a nonlinear system is difficult in general, although it may be possible to prove the existence and the uniqueness of the equilibrium in the system. Normally a nonlinear system is linearized by taking the first order terms of Taylor expansion round the equilibrium. Then the resulting linear system is examined for the stability or the asymptotic stability of the equilibrium. The linear system can represent the original nonlinear system only when the disturbances to the equilibrium are small. Therefore, the stability or the asymptotic stability of the equilibrium in the linear system corresponds to the *local stability* or the *local asymptotic stability* of the equilibrium in the nonlinear system, although in some cases the equilibrium may be globally stable and may be proved to be so in the nonlinear system.

It should be pointed out that the local stability and the asymptotic stability described here are not the same ideas as those in the car-following model mentioned in Chapter 2. This distinction will be clarified in Chapter 4.

The main concern about stability in this thesis is the stability of equilibria. However, it is worth mentioning that the above definition of the stability of equilibria may be extended to discuss the stability of other types of steady-state solutions. Consider first a periodic-2 orbit defined by two points, $\mathbf{x}^{(1)}$ and $\mathbf{x}^{(2)}$,

in the discrete system (3.2) such that

$$\mathbf{x}^{(1)} = \mathbf{f}(\mathbf{x}^{(2)}), \text{ and } \mathbf{x}^{(2)} = \mathbf{f}(\mathbf{x}^{(1)}).$$

It can be observed that the periodic-2 orbit corresponds to a fixed point $\mathbf{x}^{(1)}$ (or equivalently $\mathbf{x}^{(2)}$) in the system

$$\mathbf{x}(n+2) = \mathbf{f}(\mathbf{f}(\mathbf{x}(n))) \equiv \mathbf{f}^{(2)}(\mathbf{x}(n))$$

with

$$\mathbf{x}^{(1)} = \mathbf{f}^{(2)}(\mathbf{x}^{(1)}), \text{ or } \mathbf{x}^{(2)} = \mathbf{f}^{(2)}(\mathbf{x}^{(2)}).$$

Therefore, the stability of the fixed point in the second system implies that of the periodic orbit in the original system (3.2). Similarly, the stability of a period-K orbit in (3.2) is equivalent to that of a fixed point in the system

$$\mathbf{x}(n+K) = \mathbf{f}(...(\mathbf{f}(\mathbf{x}(n)))...) \equiv \mathbf{f}^{(K)}(\mathbf{x}(n)),$$

with the fixed point $\mathbf{x}^*$ given by

$$\mathbf{x}^* = \mathbf{f}(...(\mathbf{f}(\mathbf{x}^*))...) \equiv \mathbf{f}^{(K)}(\mathbf{x}^*). \tag{3.6}$$

Note that $\mathbf{x}^*$ can be any one of the K points on the periodic orbit.

Considerations of the stability of a periodic orbit in the continuous system (3.1) need a little more treatment. A useful technique to examine the phase space of a continuous system is to consider the intersections of a trajectory with a suitably chosen hyperplane, or the *Poincaré section.* Imagine a trajectory in a three-dimensional phase space intersected by a two-dimensional plane. The trajectory may hit the plane repeatedly (though not necessarily with equal time intervals) and so may produce a sequence of intersecting points. In this way, the continuous system is converted to a discrete system represented by a mapping of the two-dimensional plane to itself and the continuous trajectory replaced by the set of intersecting points. Clearly, a periodic orbit in the continuous system may be replaced by one point or a finite set of points and so corresponds respectively

to a fixed point or a periodic orbit in the discrete system. Thus the stability of the fixed point or periodic orbit in the discrete system reflects that of the periodic orbit in the continuous system. To examine the stability of a periodic orbit in the continuous system, we can simply inspect the intersections of adjacent trajectories in the continuous system with a Poincaré section and consider the stability of the resulting fixed point or periodic orbit in the discrete system. The idea is the same in the case of higher dimensional systems, although an appropriate hyperplane needs to be employed. More details about the technique can be found in Parker and Chua (1989). The analysis of the stability of a quasi-periodic or chaotic orbit is much more difficult and is outside the area of this thesis.

Attracting steady-state solutions are of practical interest since unstable steady-state solutions do not occur naturally, that is, when the system is running in practice or in computer simulations. Unless starting from the solution itself exactly, the trajectories usually diverge from an unstable solution.

3.1.4. Bifurcations and bifurcation diagrams

Normally any dynamical system contains at least one parameter. Parameters have values which are fixed in any particular application of a dynamical system, but which may be different in other applications. For example, the reaction time τ, the constants l and m are parameters in the car-following model (2.1). If the values of parameters are changed, the steady state of the system may also change. For a small change in the value of a parameter, there may be only a small quantitative change in the steady state, such as a small change in the position, shape, or size of an attractor. The steady state may also experience qualitative changes as the parameter value changes. For example, an attractor may lose its stability, or a new attractor may begin to emerge. Such a qualitative change is called a *bifurcation*. The parameter concerned is called the *control parameter*, or the *bifurcation parameter*. The value at which a bifurcation takes place is the *bifurcation value*.

As the value of a parameter varies, the steady state of a system may undergo a series of bifurcations. This bifurcation sequence may be best shown by a *bifurcation diagram*, which is a plot of the steady states against the values of a parameter. Figure 3.4a shows the bifurcation diagram for the logistic map

(Thompson and Stewart, 1988)

$$x(n+1) = \mu\, x(n)\, (1-x(n)),$$

where μ is the parameter considered. For a given value of the parameter μ, one point in the diagram means a stable fixed point solution, two points mean a stable period–2 orbit, and so on. A chaotic attractor is signified by infinite numbers of points. Note that only steady states are plotted. When μ is less than 3, the system has a stable fixed point. When μ is 3, the fixed point bifurcates into a period–2 orbit. As μ is increased gradually, this period–2 orbit becomes a period–4 orbit, which then turns into a period–8 orbit. The process continues until the steady states become irregular. This sequence in which the period doubles itself as the parameter varies is called *period doubling* and is a typical *route to chaos* (Thompson and Stewart, 1988). Here, it can be seen that chaos begin to occur when μ is between 3.5 and 3.6. Another feature in this diagram is that there are some periodic solutions in the regime of chaos. These are the *periodic windows*. To show that the irregular behaviour is chaotic, Liapunov exponents are plotted against μ in Figure 3.4b, where it can be seen that they are negative for non-chaotic attractors and positive in the regime of chaos. Where there are periodic windows, the exponents become negative.

Bifurcation diagrams require a large amount of computer time and memory to produce. The logistic map is very simple. Models such as the gravity model and the trip assignment model are much more complicated.

3.2. Methods for studying nonlinear dynamical systems

Analysis of a dynamical system normally involves identifying possible steady states of the system and finding out how these states and their stability change with the values of parameters, so as to get a general view of the dominant, if not complete, dynamic behaviour of the system. Particular attractors of interests may be examined further. On the one hand, it may be desirable to investigate the physical (practical) meaning of a point attractor. It is not very unusual for a system to have more than one point attractor; one of them may be better than the other from the practical point of view. On the other hand, it may also be desirable to examine chaotic behaviour if there is evidence that it exists in a system. As mentioned above, not all dynamical systems are chaotic. When an

attractor is apparently chaotic, it should be characterized by Liapunov exponents and fractal dimensions to show that it *is* chaotic. Meanwhile, research in chaos is still developing. It is interesting to apply the new concepts and techniques to investigate chaotic behaviour.

A dynamical system modelled by differential or difference equations may be analysed both theoretically and numerically. As mentioned above, dynamical systems such as (3.1) and (3.2) generally cannot be solved analytically. However, it may be possible to investigate the existence, the uniqueness, and the stability of an equilibrium without solving the equations. Methods used in theoretical analysis are often different for different models and they will be described when they are met in the thesis.

Numerical analysis can be used to deal with more complicated situations where theoretical analysis is not possible. For example, it can locate and examine any kind of attractors, particularly the chaotic ones. The main task of numerical analysis of a dynamical system is to find all possible attractors for various initial conditions and values of parameters. This is a daunting task, especially for multi-dimensional models with several parameters. Clearly, it is not possible to exhaust *all* possible initial conditions, values of parameters, and their various combinations in the system. In practice, however, the main concern is the typical or the dominant behaviour that prevails in the system for most of the time. Numerical calculations can be made to find these dominant forms of behaviour in the system. Methods for numerical calculations are described below.

3.2.1. Finding steady-state solutions

Steady-state solutions can be found by solving the evolutionary equations numerically (integrating differential equations for continuous systems or iterating difference equations for discrete systems) for given initial conditions and values of parameters until the system reaches a steady state. This method has the advantage of being straightforward to program and of being able to find any kind of attractor. It also has some drawbacks. First, the transients may be very long so that it may take time for the system to reach a steady state. Secondly, the method can only find stable solutions; it fails for unstable steady-state solutions since they are repelling. Thirdly, it is difficult, if not impossible, for a program

(and often the user) to tell when a steady state has been achieved. A manual monitoring is often needed, making the computing inefficient.

Another way to find a steady-state solution is by use of a zero finding algorithm. For example, an equilibrium can be found by solving the nonlinear equations (3.3) or (3.4) by a zero finding algorithm. In principle, one can also use a zero finding algorithm to find a periodic orbit by applying the algorithm to (3.6) in the discrete case, or to a Poincaré section in the continuous case. See Parker and Chua (1989) for more details. The advantage of the zero finding method is that it is quite efficient and can find both stable and unstable equilibria. However, it is subject to the usual drawback of any zero-finding algorithm: the initial guess must be sufficiently close to the actual value.

The three traffic models to be considered are all very complicated. The zero-finding method is less feasible. In addition, it is the stable steady-state solutions that are of greatest interest. Therefore, the integration or iteration method is used in most cases. The zero-finding method is used only as an assistance in the bifurcation analysis to identify unstable equilibrium points when their existences and approximate values are indicated.

3.2.2. Examination of particular attractors

Once an attractor is found, it may be viewed by time series plots and phase portrait plots. An equilibrium may be recognized without any plot. On the other hand, there may be situations where it is difficult to tell if an attractor is periodic, quasi-periodic, or chaotic by simply viewing the plots. Spectral analysis may be used to determine the periodicity of a time series. The power spectrum of a periodic solution contains spikes at integer multiples of the *fundamental frequency* of the solution; the power spectrum of a quasi-periodic solution consists of spikes at integer multiplies of the *frequencies which are various sum and differences of the set of base frequencies* of the solution; the power spectrum of a chaotic attractor is a broad band, though there may be spikes. In principle, it is possible to distinguish periodic and quasi-periodic solutions by spectral analysis because the spectrum for the former has spikes at integer multiples of one particular frequency while the spectrum for the latter has spikes at integer multiples of more than one frequency. In other words, the peaks of the spectrum of a quasi-periodic solution may not be spaced at frequencies whose ratios are

rational. In practice, however, it is hardly possible to differentiate the two types of solution by spectral analysis because of the difficulty in determining whether two particular frequencies in a spectrum have a rational or irrational ratio. The usefulness of spectral analysis is that it can indicate stochastic behaviour or chaos by a broad-band, continuous spectrum. In this thesis, spectral analysis is used to test the periodicity when the time series of an attractor appears to be irregular. Liapunov exponents and fractal dimensions are calculated for the attractor if the spectral analysis indicates that the attractor is chaotic.

3.2.3. Calculation of bifurcation diagrams

Calculation of bifurcation diagrams involves increasing the value of a control parameter step by step and finding the steady state of the system at each step by integrations or iterations. The integrations or iterations are made in two stages. In the first stage, the system is integrated or iterated for a period of time so that transients have decayed. In the second stage, the system is integrated or iterated for another period of time; the solution is assumed to be the steady state and is plotted on the bifurcation diagram. To ensure that a steady state is reached, the calculations in the first stage must be very long, making the calculation very time consuming. Initial conditions may be taken to be the same for all steps of parameters. Alternatively, the final state for the previous parameter value can be used as the initial condition for the current step, leaving only the initial condition for the first step to be specified. In this way, the length of transients may be reduced. In this thesis, both methods will be used in choosing the initial conditions for producing the bifurcation diagrams for the gravity model. Different steady states and so different bifurcation sequences may be found by using different initial conditions.

3.2.4. Analyses in this thesis

Given a dynamic traffic model, the research in this thesis tries to answer the following questions: Is there an equilibrium in the model? If there is one, is it stable? What other types of attractor does the model have? Is there chaos in the model? Equilibria are examined theoretically wherever possible; the methods used may differ for the three models. Other types of dynamic behaviour are investigated numerically. The analysis to be made for the three models are

outlined below.

The car-following model

The stability of an equilibrium in a linear car-following model is examined theoretically by analyzing the roots of the characteristic equation of the model; an equilibrium in a nonlinear model is examined by linearization. Numerical analysis involves integration of the car-following equations. There is not a standard algorithm for integrating delay-differential equations, so the Runge-Kutta algorithm for integrating ordinary differential equations has been modified to deal with the car-following model. A program was developed to integrate the car-following equations, and the program is used to identify possible attractors in the model for various initial conditions and parameters values.

The gravity model

The existence, the uniqueness, and the stability of an equilibrium is investigated theoretically. The practical implications of an equilibrium are explored. Numerical calculations are made by iterating the model. The initial conditions and the values of parameters are "scanned" so as to find almost all possible attractors. The calculations were often made in a batch mode on a computer because the amount of computation is very large.

The logit-based trip assignment model

This model is mathematically similar to the gravity model with the exponential deterrence function. Therefore, it is examined in similar way to that for the gravity model.

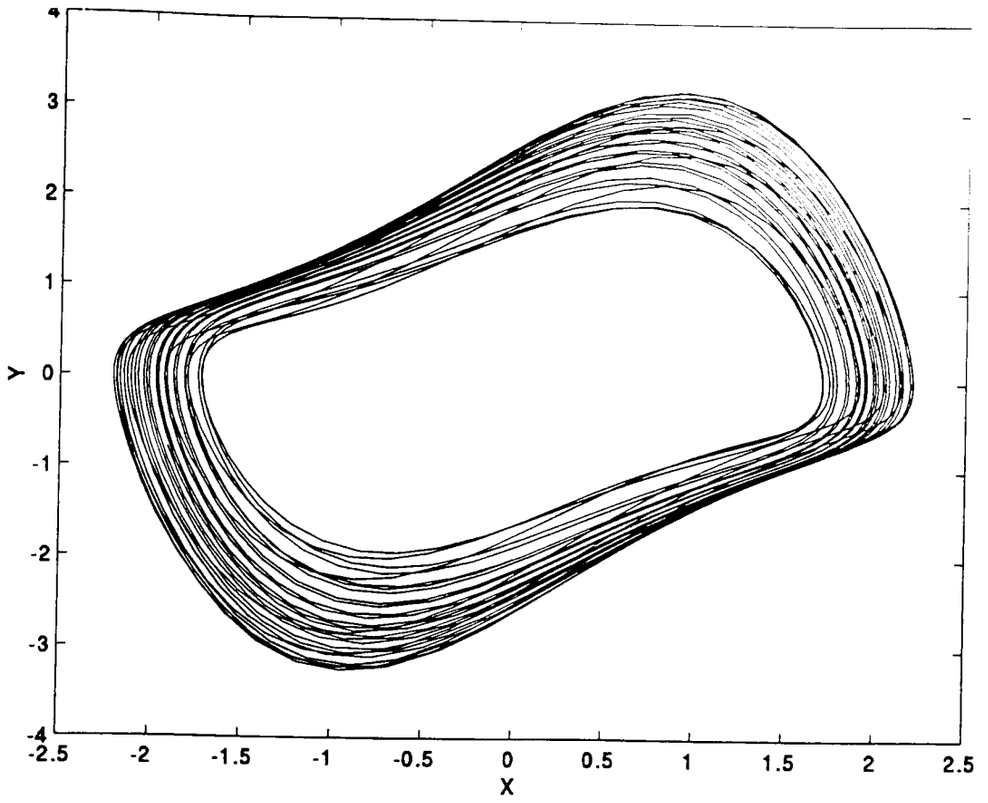
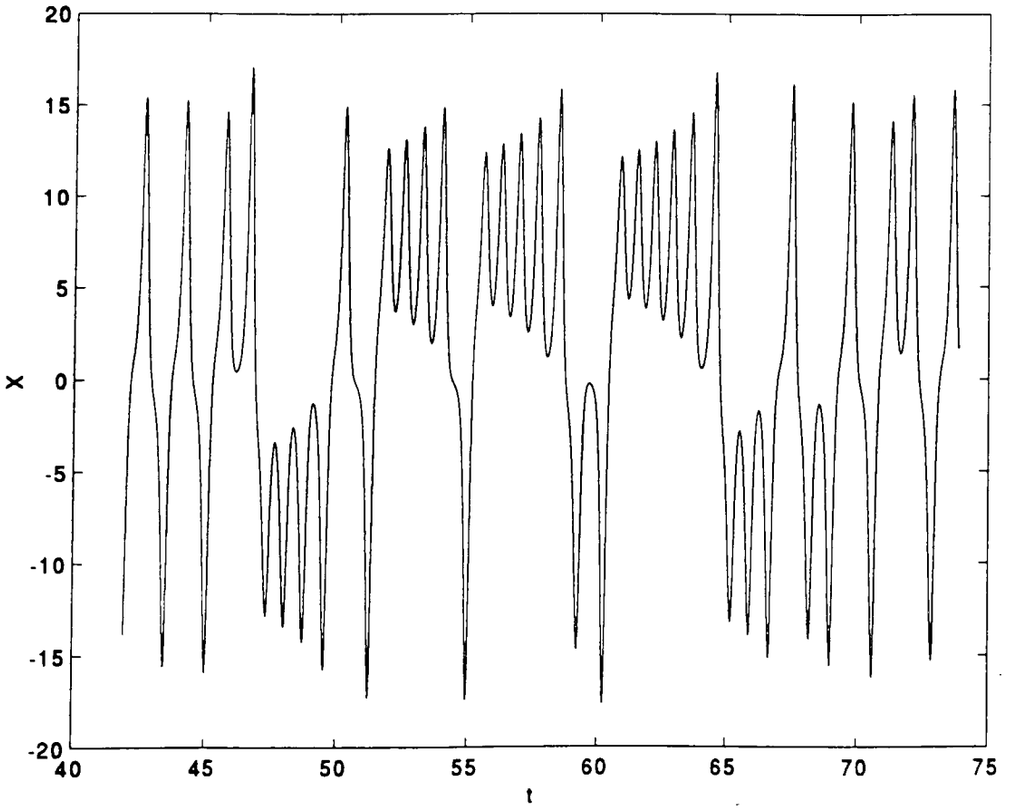
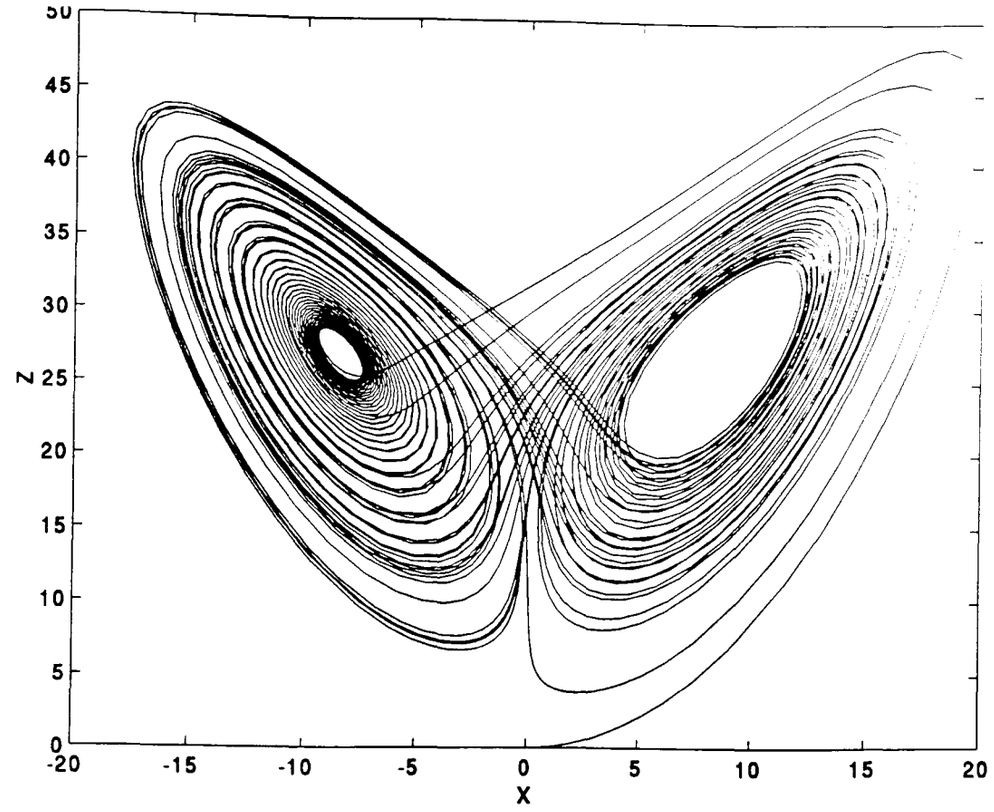


Figure 3.1 Phase portrait projection of a quasi-periodic solution of the forced van der Pol equation.

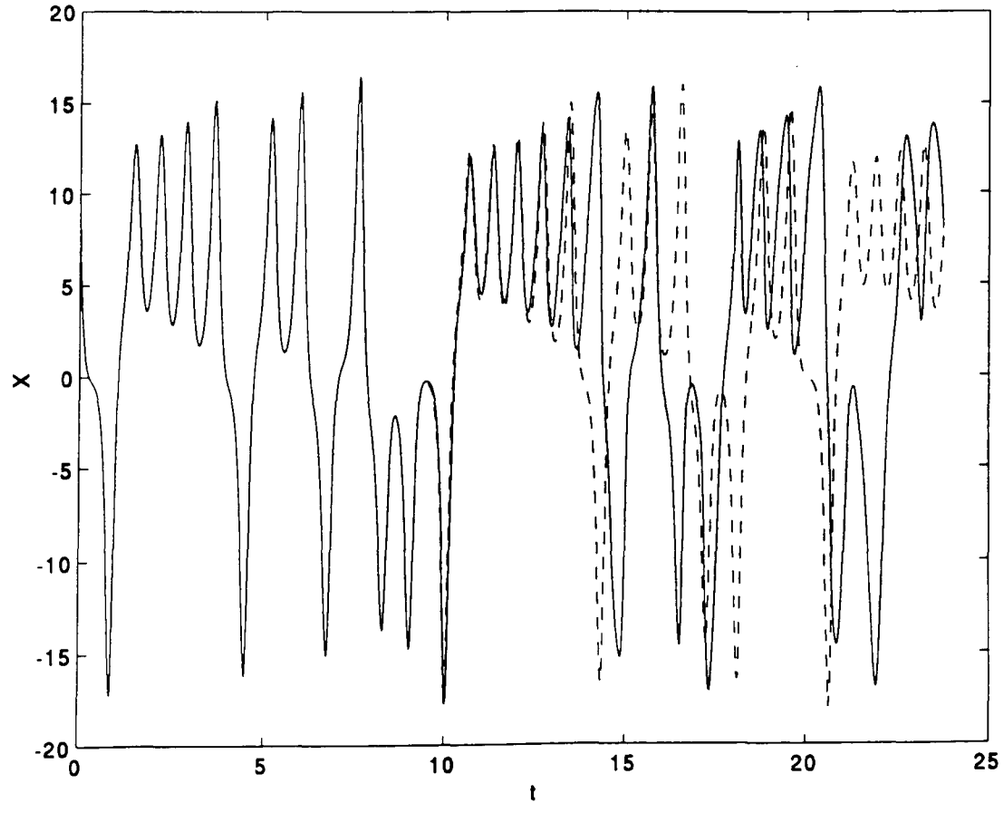


(a)

Figure 3.2 Chaotic attractor of the Lorenz equations. (a) Time series.

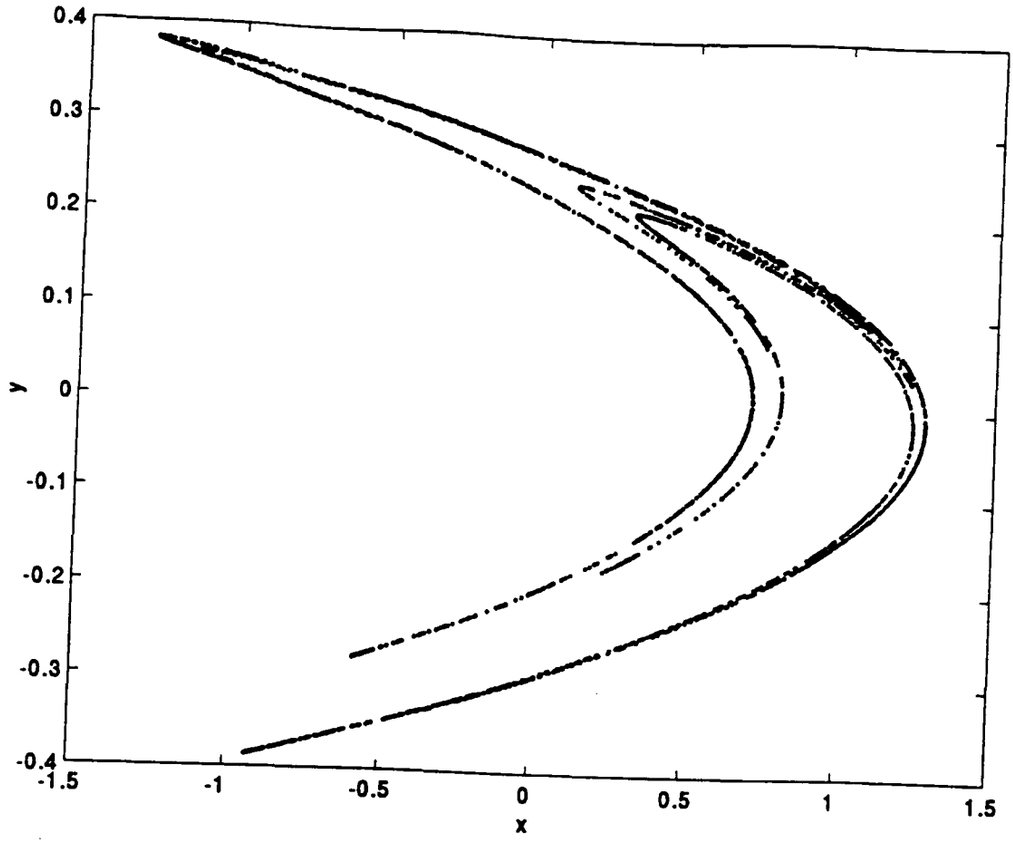


(b)

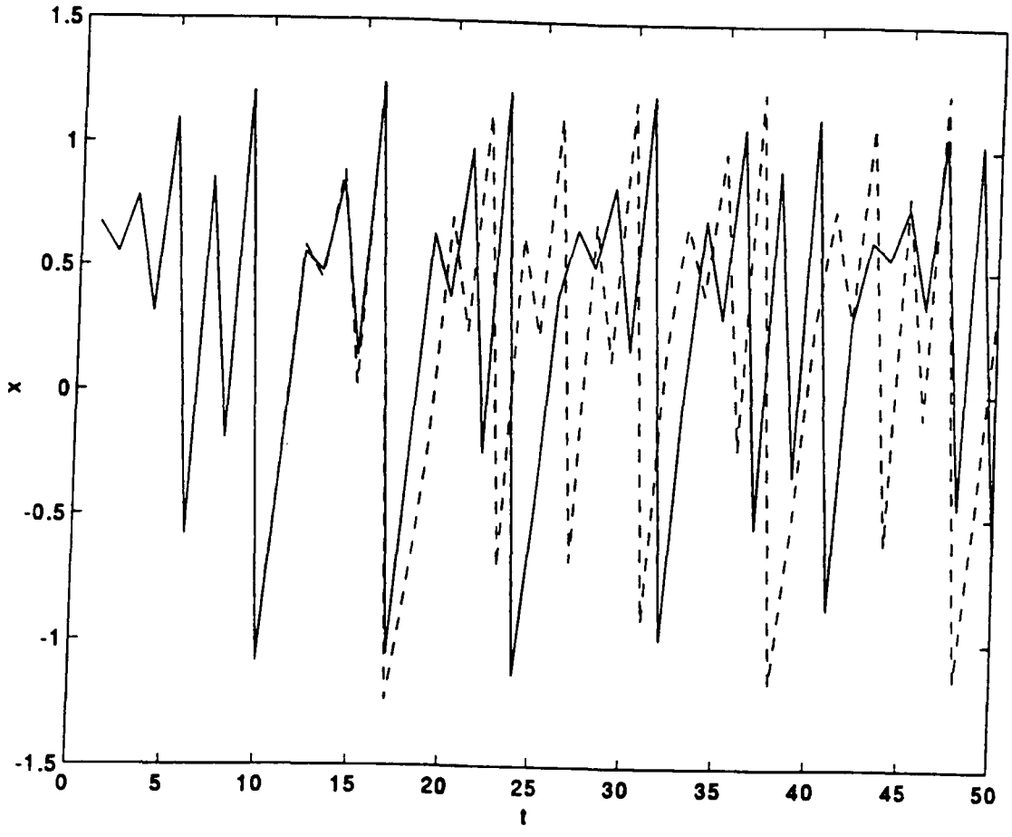


(c)

Figure 3.2 continued. (b) Trajectory in (x,z)-plane; (c) sensitivity to initial conditions.

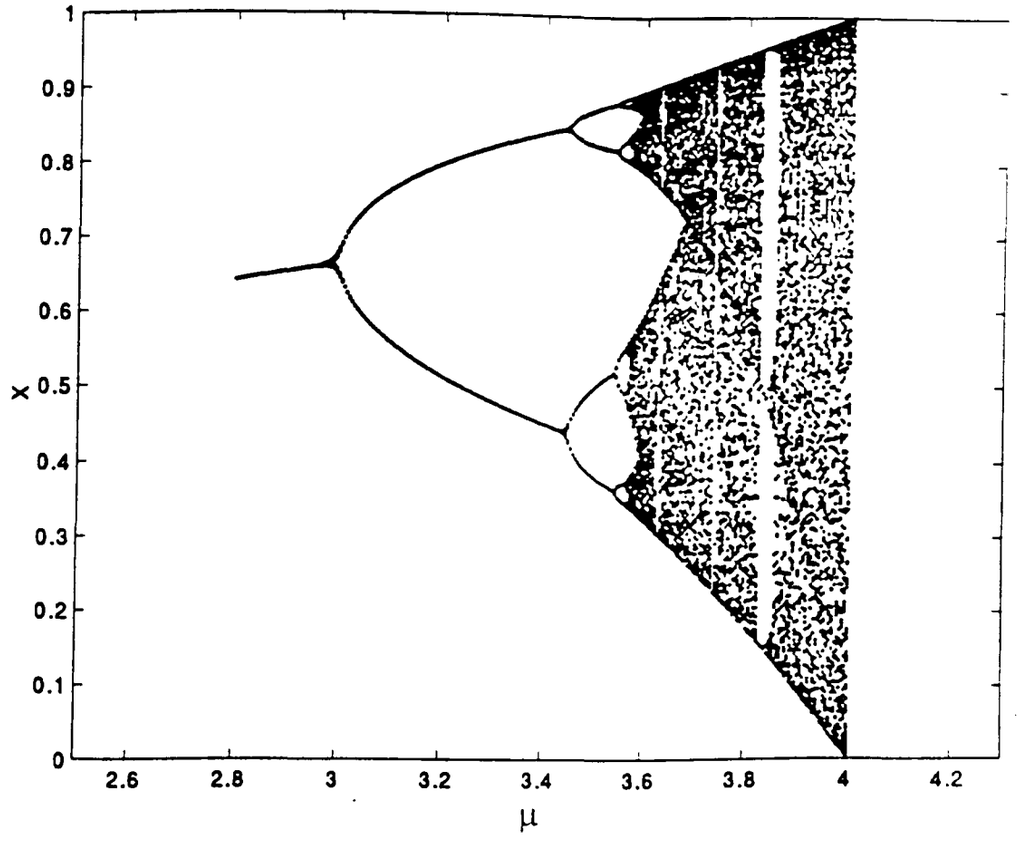


(a)

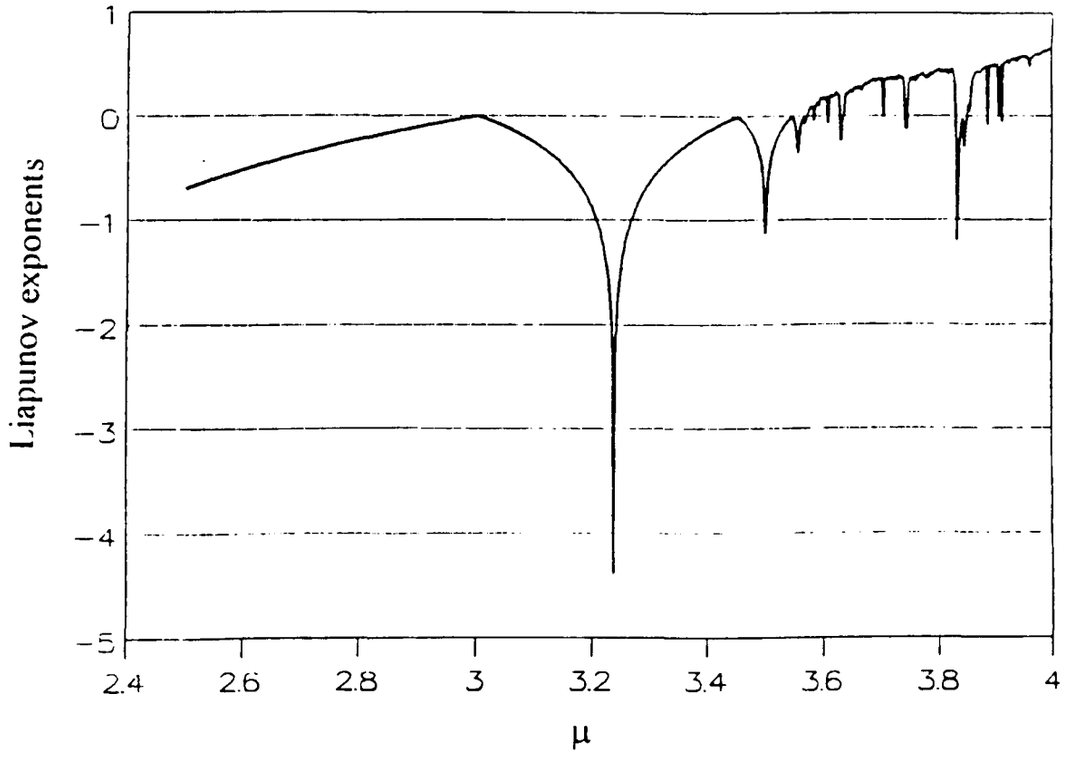


(b)

Figure 3.3 Chaotic attractor of the Hénon map. (a) Phase portrait; (b) sensitivity to initial conditions.



(a)



(b)

Figure 3.4 Bifurcation of the logistic map. (a) Bifurcation diagram; (b) Liapunov exponent.

CHAPTER 4. THE DYNAMIC BEHAVIOUR OF THE CAR-FOLLOWING MODEL

In this chapter the dynamic behaviour of the movement of traffic flow on a road link is investigated based on the car-following model. The investigation aims to find out how the speed of each car and the spacing between the cars vary with time, and how disturbances such as accelerations or decelerations of a leading car affect the cars behind.

4.1. INTRODUCTION

The car-following model (2.1) consists of a system of delay-differential equations modelling a line of cars following each other on a single lane of a stretch of road. It is assumed that a driver responds, through accelerations or decelerations and with a reaction time, to the relative speed to the car in front. Let $x_n(t)$ be the position of nth car in the line at time t. Then the car-following equations can be written as

$$\ddot{x}_n(t) = \beta_n(t)\ (\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)), \quad n = 2, 3, \cdots, N, \tag{4.1}$$

where $\beta_n(t)$ is the sensitivity of car n and is given by

$$\beta_n(t) = \frac{\alpha\ (\dot{x}_n(t))^m}{(x_{n-1}(t-\tau) - x_n(t-\tau))^l}, \quad n = 2, 3, \cdots, N.$$

In these equations, dots denote time derivatives. τ is the reaction time. α is a (positive) coefficient of proportionality; m and l are non-negative constants, not necessarily integers. The motion of the first car, defined by $\ddot{x}_1(t)$ and $\dot{x}_1(t)$, is normally taken to be given. It can be considered as a forcing term.

When $\ddot{x}_1 \equiv 0$ and $\dot{x}_1 \equiv$ constant, the model is autonomous, otherwise it is non-autonomous. On the other hand, if $m = l = 0$, the model is linear and the sensitivity β_n reduces to α, otherwise the model is nonlinear. Consequently, we may have four types of model: linear autonomous model, linear non-autonomous model, nonlinear autonomous model, and nonlinear non-autonomous model.

Although it is a system of simultaneous equations, the car-following model has a special feature: each equation contains only the variables (speed and position) of two adjacent cars; the output of each equation can be considered as the input of the next equation. Therefore, given the motion of the first car, the equations can often be dealt with one by one successively both in theoretical analysis and in numerical studies.

The car-following model (4.1) is a system of second-order delay-differential equations, which can always be converted to a system of first-order equations by introducing extra variables $u_n(t) = \dot{x}_n(t)$, $n = 2, 3, \cdots, N$. Therefore, the state variables of the model are x_n, and $\dot{x}_n$, $n = 2, 3, \cdots, N$, and the solution the position and speed of cars as functions of time. In general, the solution is unbounded since x_n will usually go to infinity as t tends to infinity. Consequently, the theory and the methods of nonlinear dynamics described in Chapter 3 cannot be applied to examine the model. This is because the unbounded solutions cannot tend towards an attractor, which must always be bounded.

This difficulty can be overcome by re-expressing the car-following model in terms of relative quantities, or the relative spacing and the relative speed between cars, so that the solutions can remain bounded. Denote the spacing between adjacent cars by

$$y_n(t) + b = x_{n-1}(t) - x_n(t), \quad n = 2, 3, \cdots, N,$$

where b can be considered as a desired headway that drivers would follow under steady state or equilibrium conditions. See Figure 4.1. Then

$$\dot{y}_n(t) = \dot{x}_{n-1}(t) - \dot{x}_n(t), \quad \text{and} \quad \ddot{y}_n(t) = \ddot{x}_{n-1}(t) - \ddot{x}_n(t).$$

Denoting the relative speed $\dot{y}_n(t)$ by $v_n(t)$, the car-following model (4.1) becomes

$$\ddot{x}_n(t) = \beta_n(t)\, v_n(t-\tau), \quad n = 2, 3, \cdots, N,$$

where

$$\beta_n(t) = \frac{\alpha\, (\dot{x}_n(t))^m}{(y_n(t-\tau) + b)^l}, \quad n = 2, 3, \ldots, N.$$

Writing the equations in terms of $v_n(t)$ and $y_n(t)$ gives

$$\dot{v}_2(t) = \ddot{x}_1(t) - \beta_2(t)\, v_2(t-\tau), \tag{4.2a}$$

$$\dot{v}_n(t) = \beta_{n-1}(t)\, v_{n-1}(t-\tau) - \beta_n(t)\, v_n(t-\tau), \quad n = 3, 4, \ldots, N, \tag{4.2b}$$

$$\dot{y}_n(t) = v_n(t), \quad n = 2, 3, \ldots, N, \tag{4.2c}$$

where

$$\beta_n(t) = \frac{\alpha\, (\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau) + b)^l}, \quad n = 2, 3, \ldots, N.$$

A solution of the car-following model (4.2), $v_n(t)$ and $y_n(t)$, $n = 2, 3, \ldots, N$, should satisfy the following essential conditions to make sense:

(1) The spacing between cars should be positive to avoid collision, that is, $y_n(t) + b > 0$. In fact, at least for $l > 0$, the solutions break down if $y_n(t) + b$ becomes 0. Although in practice it is normally required that $y_n(t)$ is non-negative for safety reasons, the first condition will be used in numerical analysis in order to find all possible types of behaviour in the model.

(2) The spacing between cars should also be small enough to guarantee that the car-following model is applicable. When traffic flow is sparse the spacing between cars may be so big that the motion of the leading car has little effect on the following car. It can be seen from the car-following equations that for $l > 0$ the sensitivity and so the accelerations of each

car will tend to zero if $y_n(t) + b$ goes to infinity.

(3) The speed of each car should be non-negative. The speed of the leading car is assumed non-negative, so the reverse motion of the following car is considered to be unreasonable.

(4) The speed of each car should also be bounded for obvious reasons. Thus the relative speed between the cars is bounded as well.

Using the car-following model (4.2), we can consider steady states such as equilibria and periodic motions. For example, in the autonomous system where $\ddot{x}_1 \equiv 0$ and $\dot{x}_1 \equiv$ constant, we can find an equilibrium solution by setting

$$\dot{v}_n(t) = 0, \text{ and } \dot{y}_n(t) = 0, \quad n = 2, 3, \ldots, N$$

in the equations (See Chapter 3). From this we get the equilibria

$$v_n^e = 0, \; y_n^e = a_n, \quad n = 2, \ldots, N,$$

where the a_n are constants, and the superscripts "e" denote equilibrium. The cars all move at the same speed, keeping the relative spacing constant. Such an equilibrium exists for any set of values $a_n \geq 0$: there is a continuum of equilibria.

Stability of a steady-state solution is an immediate concern in a dynamical system once the solution is found. The stability concepts used in the literature of the car-following model are a little different from those in nonlinear dynamics. As mentioned in Chapter 2, two types of stability in the car-following model have been discussed in the literature, namely, local stability and asymptotic stability (Chandler *et al.*, 1958 and Herman *et al.*, 1959). Local stability was defined for the equilibrium solution where the speed of each car is equal and the spacing between the cars constant. The equilibrium is (locally) stable if disturbances in the speeds of cars are damped with time. Asymptotic stability, on the other hand, is concerned with whether or not the disturbances in the acceleration of the first car are damped along the line of cars.

The concept of stability, including both local stability and asymptotic stability,

of equilibria in a general (nonlinear) dynamical system has been described in Chapter 3. This description is repeated here very briefly. An equilibrium is stable if all trajectories starting near the equilibrium point stay nearby for all time. It is asymptotically stable if all nearby trajectories approach the equilibrium point as time tends to infinity. For nonlinear systems, these are both local forms of stability: solutions not starting sufficiently close to a stable equilibrium may move away from it.

Clearly, the local stability in the car-following model is similar to the asymptotic stability in nonlinear dynamics in that they both deal with time variations of disturbances, although the former is less general than the latter. The asymptotic stability in the car-following model, however, is essentially different from either concept in nonlinear dynamics.

In this chapter, the terms local stability and asymptotic stability in the literature of car-following model will be abandoned and those in nonlinear dynamics used to characterize the time variations of small disturbances for the car-following model. The stability of equilibria in the linear model will be investigated; the local stability of equilibria in the nonlinear model will be examined by linearization. The variations of disturbances along the line of cars will also be considered and will be referred to as *stability over cars*. A line of cars is said to be *stable over cars* if small fluctuations in the movements of the first car are not amplified as they pass from one car to the next in the line. Otherwise, if the fluctuations are amplified, then the system is unstable.

In studies of the car-following model in the literature, the coefficient of proportionality α and the reaction time τ have been assumed to be the same for all drivers in the line. In this thesis, the situation where different drivers can have different α and τ will be considered. If α_n and τ_n are the coefficient of proportionality and reaction time, respectively, for the nth driver, then the car-following model can be written as

$$\dot{v}_2(t) = \ddot{x}_1(t) - \beta_2(t)\, v_2(t-\tau_2), \tag{4.3a}$$

$$\dot{v}_n(t) = \beta_{n-1}(t)\, v_{n-1}(t-\tau_{n-1}) - \beta_n(t)\, v_n(t-\tau_n), \quad n = 3, 4, \ldots, N, \tag{4.3b}$$

$$\dot{y}_n(t) = v_n(t), \quad n = 2, 3, \ldots, N, \tag{4.3c}$$

where

$$\beta_n(t) = \frac{a_n(\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau_n) + b)^l}, \quad n = 2, 3, \ldots, N,$$

$$v_n(t-\tau_n) = \dot{x}_{n-1}(t-\tau_n) - \dot{x}_n(t-\tau_n), \quad n = 2, 3, \cdots, N,$$

$$y_n(t-\tau_n) = x_{n-1}(t-\tau_n) - x_n(t-\tau_n) - b, \quad n = 2, 3, \cdots, N.$$

In the next section, theoretical analyses are made of the linear model, both autonomous and non-autonomous, and the nonlinear autonomous model to investigate the stability of the equilibria and the stability over cars. The nonlinear non-autonomous model is more difficult to analyze theoretically and numerical methods have to be employed. The nonlinear model is studied numerically in section 4.3 to identify possible types of attractors in the model. The results obtained are summarized in the last section.

4.2. THEORETICAL ANALYSIS

The car-following model is a system of delay-differential equations. The theoretical analysis of such equations is difficult. In general, they cannot be solved analytically, not even the simplest delay-differential equation such as a first order, linear, autonomous equation. It must be pointed out that the theory of delay-differential equations is not just a simple extension of the theory of ordinary differential equations, although there are some similarities. The initial conditions for ordinary differential equations are initial values of the variables at one time instant, while the initial conditions for delay-differential equations are initial *functions* over the interval of delay time. As a result, delay-differential equations are infinite dimensional systems. However, finite-dimensional attractors can exist for systems of delay-differential equations, and some such systems are believed to possess finite-dimensional strange attractors (See, for example, Farmer, 1982).

One possible way to deal with delay-differential equation is the *method of steps* (Driver, 1977). This can be illustrated by a simple example. Consider a linear, autonomous 2-car model

$$\dot{v}_2(t) = -a\, v_2(t-\tau), \quad t > 0,$$

with an initial condition

$$v(t) = \theta(t), \qquad -\tau \le t \le 0.$$

It has been assumed that the relative speed over $-\tau \le t \le 0$ is $\theta(t)$. To make the problem simple here, let $\theta(t)$ be some constant θ_0. With this constant initial condition, it is easy to solve the equation on $0 \le t \le \tau$. For there the equation becomes

$$\dot{v}_2(t) = -a\, \theta_0, \qquad 0 \le t \le \tau$$

with

$$v_2(0) = \theta_0.$$

The solution to this equation is

$$v_2(t) = \theta_0 - a\, \theta_0\, t, \quad 0 \le t \le \tau.$$

Consider, next, the interval $\tau \le t \le 2\tau$ where the equation becomes

$$\dot{v}_2(t) = -a\, \theta_0 + a^2 \theta_0 (t-\tau), \quad \tau \le t \le 2\tau$$

with

$$v_2(\tau) = \theta_0 - a\, \theta_0\, \tau.$$

The solution now is

$$v_2(t) = \theta_0 - a\, \theta_0\, t + a^2 \theta_0 (t-\tau)^2/2, \quad \tau \le t \le 2\tau.$$

We can then consider the interval $2\tau \le t \le 3\tau$, etc. This procedure can, in principle, be continued as long as desired. It gives an explicit solution up to any finite time t. In this particular case, we get a polynomial solution of degree k for $(k-1)\tau \le t \le k\tau$. In general, however, the solution will quickly become accumulated and complicated; it is difficult to draw any general conclusion about the solution. For example, if the solutions are bounded, or oscillate, or if they tend to some constant as $t \to \infty$. It is these qualitative properties of the steady-

state solutions that is of interest here. In the following subsections, the linear autonomous model, the linear non-autonomous model, and the nonlinear autonomous model are investigated in turn to identify possible steady-state solutions and their stabilities both with time and over cars.

4.2.1. The linear autonomous model

As mentioned in chapter 2, Herman *et al.* (1959) investigated the stability of a linear car-following model using Laplace transform and numerical analysis of the inversions of the Laplace transforms. The parameters α and τ were assumed to be the same for all drivers in the line of cars. Their conclusions have been described in Chapter 2. Here, the stability of the equilibria in the linear autonomous model will be examined by theoretical analysis of the characteristic equation of the model and α and τ are allowed to be different for different drivers. The theoretical basis used in this subsection comes from Györi and Ladas (1991) and Kuang (1993).

In (4.3), let $m = l = 0$, and $\ddot{x}_1 \equiv 0$. Then a linear autonomous model is

$$\dot{v}_2(t) = -\alpha_2 v_2(t-\tau_2), \tag{4.4a}$$

$$\dot{v}_n(t) = \alpha_{n-1} v_{n-1}(t-\tau_{n-1}) - \alpha_n v_n(t-\tau_n), \quad n = 3, 4, \ldots, N, \tag{4.4b}$$

$$\dot{y}_n(t) = v_n(t), \quad n = 2, 3, \ldots, N. \tag{4.4c}$$

It can be seen that (4.4a) and (4.4b) can be dealt with independently of (4.4c) which is just an integration. We will consider first (4.4a) and (4.4b), and then deal with (4.4c). But first of all, we consider (4.4a) alone. Omitting the subscripts on α and τ for simplicity, (4.4a) becomes

$$\dot{v}_2(t) = -\alpha v_2(t-\tau), \tag{4.5}$$

which is the same as the equation considered above when the method of steps is introduced. There is an equilibrium $v_2^e = 0$ in this equation. We will examine its stability and other possible solutions.

If we seek an "exponential" solution of the equation in the form $v_2(t) = e^{\lambda t}$, we are led to the following characteristic equation

$$\lambda + \alpha e^{-\tau\lambda} = 0. \tag{4.6}$$

Suppose $\lambda = \mu + i\omega$ (where μ and ω are real) is a root of the characteristic equation, then

$$v_2(t) = e^{(\mu+i\omega)t}$$

is a solution of (4.5). And both

$$v_2(t) = e^{\mu t} \cos \omega t \quad \text{and} \quad v_2(t) = e^{\mu t} \sin \omega t$$

are real solutions of (4.5). By the principle of superposition, which is valid for any linear systems including delay-differential equations, linear combinations of the exponential solutions are also solutions of (4.5). The problem is that the characteristic equation (4.6) has, in general, infinitely many (complex) roots (Driver, 1977), and cannot be solved analytically. However, the qualitative properties of the solutions of (4.5) and those of the roots of (4.6) are related and both depend on values of α and τ. First, it is known (Györi and Ladas, 1991) that the solutions of linear delay-differential equations are exponentially bounded. In other words, if Real $\lambda < \mu_0$ for every root of the characteristic equation, then there exists a constant $M > 0$ such that

$$|v_2(t)| \leq M e^{\mu_0 t}.$$

It follows immediately that if no root of the characteristic equation has a positive real part then $v_2(t)$ will tend to zero as $t \to \infty$ and the zero solution or the equilibrium of (4.5) is asymptotically stable. The equilibrium is unstable if any root has a positive real part. In this case, the corresponding solution of (4.5) tends to infinity as $t \to \infty$.

Secondly, according to Györi and Ladas (1991), every solution of equation (4.5) oscillates if and only if the characteristic equation has no real roots. Györi and Ladas (1991) also give an explicit condition for oscillations: every solution of equation (4.5) oscillates if and only if $\alpha\tau > 1/e$.

In the following, the roots of the characteristic equation are examined both theoretically and numerically. To simplify the analysis, let $\lambda' = \lambda\tau$, which is equivalent to changing the time scale in (4.5) from t to t/τ such that (4.5) becomes

$$\dot{v}_2'(t') = -\alpha' v_2'(t'-1),$$

where $t' = t/\tau$, and $\alpha' = \alpha\tau$. Then the characteristic equation (4.6) becomes

$$\lambda' + \alpha' e^{-\lambda'} = 0. \tag{4.7}$$

For the purpose of the analysis, substitute $\lambda' = \mu' + i\omega'$ into (4.7), and set both the real and imaginary parts to be 0. Then we have

$$\mu' + \alpha' e^{-\mu'} \cos \omega' = 0 \tag{4.8a}$$

$$\omega' - \alpha' e^{-\mu'} \sin \omega' = 0 \tag{4.8b}$$

where $\mu' = \mu\tau$, and $\omega' = \omega\tau$.

First of all, from (4.8b) because $\alpha' > 0$, the root of the characteristic equation must be such that

$$\omega' \in [0, \pi)\cup(2\pi, 3\pi)\cup(4\pi, 5\pi)\cup\cdots\cup(-\pi, 0]\cup(-3\pi, -2\pi)\cup(-5\pi, -4\pi)\cup\cdots.$$

The root of the characteristic equation is the point where (4.8a) and (4.8b) intersect in the (μ',ω')-plane. It can be easily verified that the equations and roots are symmetric with respect to the μ' axis of the (μ',ω')-plane. Therefore we need initially consider only positive ω'. The intersections of the curves for α' equal to 0.5 and 50 and for ω' within $(-\pi, \pi)$, $(2\pi, 3\pi)$, and $(4\pi, 5\pi)$ are shown in Figure 4.2. It can be seen that, for the same value of α', μ' decreases if ω' is increased by 2π. This is because, from (4.8b),

$$\mu' = \ln\left(\frac{\alpha' \sin \omega'}{\omega'}\right), \quad \omega' \neq 0.$$

Increasing ω' by 2π changes μ' to

$$\ln\left(\frac{\alpha' \sin(\omega'+2\pi)}{\omega'+2\pi}\right) < \ln\left(\frac{\alpha' \sin\ \omega'}{\omega'}\right),$$

although in (4.8a) μ' is unchanged. Clearly, this feature will not change if ω' is negative and changed to $\omega'-2\pi$. Therefore, for given α', larger values of μ' appear when $|\omega'|$ is small. Thus, given α', then μ' is the biggest when $0 < |\omega'| < \pi$. It can also be seen in Figure 4.2 that as the value of α' increases, the curves and so the intersection points move from left to right, which means, at least for the values plotted here, that μ' increases with α'. In fact, this is true for all values of α' provided that $|\omega'| > 0$ because (4.8b) can also be written as

$$\mu' = \ln \alpha' + \ln\left(\frac{\sin\ \omega'}{\omega'}\right)$$

which shifts from left to right in the (μ', ω')-plane as α' increases.

We can find out more about how μ' and ω' vary with α'. When $\omega' \neq 0$, from (4.8) we get

$$\mu' = -\omega' \cot \omega', \tag{4.9a}$$

$$\alpha' = \omega' \, e^{\mu'} / \sin \omega', \tag{4.9b}$$

where $\omega' \neq \pm k\pi$, $k = 1, 2, \cdots$. While if $\omega' = 0$, then from (4.8a) we have

$$\alpha' = -\mu' \, e^{\mu'}. \tag{4.9c}$$

Thus given values of ω' we can calculate μ' and then α'. In (4.9a) and (4.9b), if $\omega' \to 0$, then $\mu' \to -1$, and $\alpha' \to 1/e$. But if $\omega' \to 2k\pi+0$ then $\mu' \to -\infty$, and $\alpha' \to 0$. While in (4.9c) α' has one maximum value $1/e$ at $\mu' = -1$, and μ' must be negative since $\alpha' > 0$. All these can be seen in Figures 4.3a and 4.3b, which show both μ' and ω' versus α' for $\omega' \in [0, \pi)$ and $\omega' \in (2\pi, 3\pi)$, respectively.

Now we can find the critical value of α' at which μ' and ω' change qualitatively. First, suppose that $\mu' > 0$, and so $e^{-\mu'} < 1$. Then we have

$$\cos \omega' < 0 \quad \text{and} \quad |\omega'| < \alpha'$$

from (4.8a) and (4.8b), respectively. Clearly, when $0 < \alpha' < \pi/2$, $\cos \omega' > 0$ and no root of positive μ' exists. The equilibrium $v_2^e = 0$ in (4.5) is therefore asymptotically stable.

Secondly, suppose that $\mu' = 0$. Then we have

$$\alpha' = |\omega'| \quad \text{and} \quad \omega' = \pm (2k\pi + \frac{\pi}{2}), \quad k = 0, 1, \cdots.$$

That is, when $\alpha' = 2k\pi + \pi/2$, there is a pair of roots $\mu' = 0$, $\omega' = \pm (2k\pi + \pi/2)$, where $k = 0, 1, \cdots$. To see how μ' changes at these points as α' varies, find the derivatives of μ' with respect to α' at $\lambda' = \pm i\omega'$ in (4.7). We obtain

$$\begin{aligned}
\frac{d\mu'}{d\alpha'}\Big|_{\lambda'=\pm i\omega'} &= \text{Real}\left[\frac{d\lambda'}{d\alpha'}\right]\Big|_{\lambda'=\pm i\omega'} \\
&= \text{Real}\left[\frac{\lambda'}{\alpha'(1 + \lambda')}\right]\Big|_{\lambda'=\pm i\omega'} \\
&= \text{Real}\left[\frac{\omega'^2 \pm i\omega'}{\alpha'(1 + \omega'^2)}\right] \\
&= \frac{\omega'^2}{\alpha'(1 + \omega'^2)} > 0.
\end{aligned}$$

This implies that the real part of all the roots increases at $\mu' = 0$ as α' increases. In other words, μ' changes from negative to positive at $\omega' = \pm\pi/2$ when $\alpha' = \pi/2$, at $\omega' = \pm 5\pi/2$ when $\alpha' = 5\pi/2$, and so on. Therefore, when $\alpha' = \alpha\tau > \pi/2$, there is at least one root of (4.7) with a positive real part and the zero solution of (4.5) is unstable. Unless we start from the equilibrium exactly, the solution tends to infinity eventually. When $\alpha\tau = \pi/2$, there is a periodic solution

$$v_2(t) = e^{\pm i\frac{\pi t}{2\tau}}$$

with a frequency $\pi/(2\tau)$. This is a *neutrally stable* period solution or a *center* : a deviation from the equilibrium in initial conditions is conserved; the system has no transient. The system bifurcates from the equilibrium at $\alpha\tau = \pi/2$; a slight

change in the value of $\alpha\tau$ can cause the solution either to tend to infinity or to approach the stable equilibrium.

Thirdly, consider under what condition there is no imaginary root in (4.8) and so there is no oscillatory solutions in the equation (4.5). Although Györi and Ladas (1991) provided the necessary and sufficient condition for all solutions of the equation (4.5) to oscillate, we are more concerned here about when the solutions do not oscillate. We have seen that if $\omega' = 0$ then

$$\alpha' = -\mu' e^{\mu'} \leq e^{-1},$$

which is a necessary condition for the solution of the equation (4.5) to be non-oscillatory. To investigate the sufficiency of the condition, we get from (4.9a) and (4.9b)

$$\alpha' = \frac{\omega'}{\sin \omega'} e^{\mu'} = \frac{\omega'}{\sin \omega'} e^{-\omega' \cot \omega'}, \quad \omega' \neq \pm k\pi, k = 0, 1, \cdots.$$

We have shown that $\alpha' \to 1/e$ when $\omega' \to 0$. If $0<|\omega'|<\pi$, then $\omega' \cot \omega' < 1$ and $\frac{\omega'}{\sin \omega'} > 1$. So we have

$$\frac{\omega'}{\sin \omega'} e^{-\omega' \cot \omega'} > 1/e, \quad 0<|\omega'|<\pi,$$

which means that, for $0<|\omega'|<\pi$, α' must be larger than e^{-1}. Thus there is no root with $\pi>|\omega'|>0$ when $\alpha' \leq e^{-1}$. It is interesting, although not necessary, to show that α' actually increases monotonically with $|\omega'|$, where $0<|\omega'|<\pi$. We find the derivative

$$\frac{d\alpha'}{d\omega'} = \frac{\sin \omega' - \omega' \cos \omega'}{\sin^2 \omega'} e^{-\omega' \cot \omega'} + \frac{\omega'}{\sin \omega'} e^{-\omega' \cot \omega'} \left(-\cot \omega' + \frac{\omega'}{\sin^2 \omega'}\right)$$

$$= \frac{\omega'}{\sin^2 \omega'} e^{-\omega' \cot \omega'} \left[\left(\frac{\sin \omega'}{\omega'} - \cos \omega'\right) + \left(\frac{\omega'}{\sin \omega'} - \cos \omega'\right)\right].$$

Since $\frac{\omega'}{\sin \omega'} > \frac{\sin \omega'}{\omega'} > \cos \omega'$ where $0<|\omega'|<\pi$, we obtain

$$\frac{d\alpha'}{d\omega'} > 0 \text{ if } \omega' \in (0, \pi), \text{ and } \frac{d\alpha'}{d\omega'} < 0 \text{ if } \omega' \in (-\pi, 0),$$

which is what we claimed. In addition, if $\omega' \to \pm\pi$ then $\mu' \to \infty$ and $\alpha' \to \infty$. Therefore $\alpha' \leq e^{-1}$ is a necessary and sufficient condition under which there is no oscillatory solution with $|\omega'| < \pi$. It can be easily verified numerically, or at least it can be seen in Figure 4.3b, that there are non-zero solutions of ω', where $\omega' > \pi$, or rather $\omega' > 2\pi$, when $\alpha' \leq 1/e$.

The above discussions about the solutions of the linear autonomous 2-car model (4.5) can be summarized as follows in terms of $\alpha\tau$ and $\omega\tau$ rather than α' and ω'.

(1) When $\alpha\tau > \pi/2$, the equilibrium $v_2^e = 0$ is unstable. The solution oscillates and its amplitude increases with time for all initial conditions except at the equilibrium itself.

(2) When $\alpha\tau = \pi/2$, there is a neutrally stable periodic solution with a frequency $\pi/(2\tau)$.

(3) When $1/e < \alpha\tau < \pi/2$, the equilibrium is asymptotically stable. The solutions are damped oscillations.

(4) When $\alpha\tau \leq 1/e$, the equilibrium is asymptotically stable. There is no oscillatory solution with angular frequency $|\omega| < \pi/\tau$, but there are oscillatory solutions of higher frequencies.

The car-following equation (4.5) is a first order, linear, homogeneous delay-differential equation with constant coefficient. The oscillating solution in the equation is a significant feature distinguishing from a first order, linear homogeneous ordinary differential equation with constant coefficient, which cannot have an oscillating solution.

The discussions of the 2-car model can be generalized to consider more than two cars. The equations (4.4a) and (4.4b) can be written in matrix form as

$$\dot{\mathbf{v}}(t) = \Sigma_{k=2}^{N} \mathbf{B}_k \mathbf{v}(t-\tau_k), \tag{4.10}$$

where

$$\dot{\mathbf{v}}(t) = (\dot{v}_2(t),\ \dot{v}_3(t),\ \ldots,\ \dot{v}_N(t))^{\mathrm{T}},$$
$$\mathbf{v}(t-\tau_k) = (v_2(t-\tau_k),\ v_3(t-\tau_k),\ \ldots,\ v_N(t-\tau_k))^{\mathrm{T}},$$

and $\mathbf{B}_k$ is a matrix with only two nonzero elements:

$$\begin{aligned} [\mathbf{B}_k]_{ij} &= -\alpha_k \qquad && i = k\text{–}1,\ j = k\text{–}1, \\ [\mathbf{B}_k]_{ij} &= \alpha_k && i = k,\ j = k\text{–}1. \end{aligned}$$

For example, a 4-car model can be written as

$$\begin{bmatrix} \dot{v}_2(t) \\ \dot{v}_3(t) \\ \dot{v}_4(t) \end{bmatrix} = \begin{bmatrix} -\alpha_2 & 0 & 0 \\ \alpha_2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_2(t-\tau_2) \\ v_3(t-\tau_2) \\ v_4(t-\tau_2) \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\alpha_3 & 0 \\ 0 & \alpha_3 & 0 \end{bmatrix} \begin{bmatrix} v_2(t-\tau_3) \\ v_3(t-\tau_3) \\ v_4(t-\tau_3) \end{bmatrix}$$
$$+ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -\alpha_4 \end{bmatrix} \begin{bmatrix} v_2(t-\tau_4) \\ v_3(t-\tau_4) \\ v_4(t-\tau_4) \end{bmatrix}.$$

The equation (4.10) has a zero solution or an equilibrium $\mathbf{v}^{\mathrm{e}} = 0$, the stability of which can be examined in the same way as that used for the 2-car model. Suppose (4.10) has a solution of the form

$$\mathbf{v}(t) = \mathbf{c}\, \mathrm{e}^{\lambda t},$$

where $\mathbf{c}$ is a constant vector, λ is a constant, both to be determined. Substituting $\mathbf{v}(t)$ into to (4.10), we have

$$\mathbf{c}\, \lambda \mathrm{e}^{\lambda t} = \mathbf{c}\, \mathrm{e}^{\lambda t} \Sigma_{k=2}^{N} \mathbf{B}_k\, \mathrm{e}^{-\lambda \tau_k}.$$

Since $\mathrm{e}^{\lambda t}$ is nonzero and can be canceled,

$$(\lambda \mathbf{E} - \Sigma_{k=2}^{N} \mathbf{B}_k\, \mathrm{e}^{-\lambda \tau_k})\, \mathbf{c} = 0,$$

where $\mathbf{E}$ is the identity matrix. For $\mathbf{c}\, \mathrm{e}^{\lambda t}$ to be a solution of (4.10), λ must be the eigenvalue of the matrix $\Sigma_{k=2}^{N} \mathbf{B}_k\, \mathrm{e}^{-\lambda \tau_k}$, and $\mathbf{c}$ the corresponding

eigenvector. For $\mathbf{c}$ to have nonzero solution, there must be

$$\det\left(\lambda\mathbf{E} - \Sigma_{k=2}^{N} \mathbf{B}_k \, e^{-\lambda\tau_k}\right) = 0. \tag{4.11}$$

This is the characteristic equation for equation (4.10). If $\lambda = \mu + i\omega$ (where μ and ω are real) is a root of the characteristic equation, and $\mathbf{c} = \mathbf{c}_{(1)} + i\mathbf{c}_{(2)}$ (where $\mathbf{c}_{(1)}$ and $\mathbf{c}_{(2)}$ are real vectors) is the corresponding eigenvector, then

$$\mathbf{v}(t) = e^{\mu t}(\mathbf{c}_{(1)}\cos\,\omega t - \mathbf{c}_{(2)}\sin\,\omega t)$$

and

$$\mathbf{v}(t) = e^{\mu t}(\mathbf{c}_{(2)}\cos\,\omega t + \mathbf{c}_{(1)}\sin\,\omega t)$$

are real solutions of (4.10). Furthermore, qualitative behaviour of the solution of the equation (4.10) depends on the roots of the characteristic equation in the same way as a single equation such as the 2-car model described above (Györi and Ladas, 1991). First, since the solutions of linear delay-differential equations are exponentially bounded, the equilibrium of the equation (4.10) is asymptotically stable if the real part of all roots of the characteristic equation is negative. Otherwise, it is unstable. Secondly, every solution of every component of $\mathbf{v}(t)$ oscillates if and only if the characteristic equation has no real roots.

To examine the root of the characteristic equation, expand (4.11) to obtain

$$(\lambda + \alpha_2 e^{-\lambda\tau_2})\,(\lambda + \alpha_3 e^{-\lambda\tau_3}), \cdots, (\lambda + \alpha_N e^{-\lambda\tau_N}) = 0,$$

or

$$\lambda_i + \alpha_i e^{-\lambda_i\tau_i} = 0, \quad i = 2, 3, \ldots, N.$$

Each of the equations has the same form as the characteristic equation (4.6) for the 2-car model. From the above analysis, it can be concluded that the equilibrium of (4.10) is asymptotically stable if

$$\alpha_i\tau_i < \pi/2, \qquad i = 2, 3, \cdots, N,$$

so that

$$\text{Real } \lambda_i < 0, \qquad i = 2, 3, \cdots, N.$$

Moreover, there is no oscillatory solution in (4.10) with a frequency lower than π if

$$\alpha_i \tau_i \leq 1/e, \quad i = 2, 3, \cdots, N,$$

although there are oscillations of higher frequencies.

Now we can deal with (4.4c), which is

$$\dot{y}_n(t) = v_n(t), \qquad n = 2, 3, \cdots, N.$$

The solution to this is simply an integration of $v_n(t)$, or

$$y_n(t) = \int_0^t v_n(s)\, ds + y_n(0), \quad n = 2, 3, \cdots, N,$$

where $y_n(0)$ is the initial condition. If $v_n(t)$ has a solution of the form $e^{\lambda t}$ then $y_n(t)$ will be of the form

$$e^{\lambda t}/\lambda + y_n(0) - (1/\lambda),$$

correspondingly. Therefore, if $v_2^e = 0$ is asymptotically stable, then $y_n(t) \equiv$ some constant is also stable. But it is not asymptotically stable because the solution $y_n(t)$ depends on the initial condition $y_n(0)$ and will become a constant whenever $v_n(t) \equiv 0$. Further, if $v_n(t)$ is a periodic solution, so is $y_n(t)$, with the same frequency but different amplitude which, again, depends on the initial condition $y_n(0)$.

The above conditions for the stability of the linear autonomous model are similar to those of Herman *et al.* (1959) listed in Chapter 2. The criteria there are given in terms of $\alpha\tau$; here they are replaced by $\alpha_i \tau_i$, $i = 2, 3, \cdots, N$, which is more general. Besides, Herman *et al.* (1959) did not consider explicitly oscillations of frequencies higher than π and so concluded that there is no oscillatory behaviour if $\alpha\tau \leq 1/e$. Here, we have shown that if $\alpha_i \tau_i \leq 1/e$, $i = 2, 3, \cdots, N$, there can be oscillations with frequency higher than π although there are none with lower frequencies.

It must be pointed out that the above stability conditions are concerned only in

the way in which the movement of each individual car varies with time. They may not guarantee the stability over cars. In other words, disturbances may be amplified along cars, even though they are damped with time. The problem of stability over cars will be considered in the next subsection.

4.2.2. The linear non-autonomous model

A linear, non-autonomous model is (putting $m = l = 0$ in (4.3))

$$\dot{v}_2(t) = \ddot{x}_1(t) - \alpha_2 v_2(t-\tau_2), \tag{4.12a}$$

$$\dot{v}_n(t) = \alpha_{n-1} v_{n-1}(t-\tau_{n-1}) - \alpha_n v_n(t-\tau_n), \quad n = 3, 4, \ldots, N, \tag{4.12b}$$

$$\dot{y}_n(t) = v_n(t), \quad n = 2, 3, \ldots, N. \tag{4.12c}$$

The stability over cars of the linear car-following model with α and τ being the same for all drivers was investigated by Chandler *et al.* (1958); their results have been described in Chapter 2. In what follows, the stability over cars is considered with α and τ being different for different drivers. The method used here is similar to that in Chandler *et al.* (1958), but the analysis is more rigorous.

To consider the stability over cars, we take the forcing term to be oscillation

$$\ddot{x}_1(t) = F \sin(\rho t - \varphi_1), \tag{4.13}$$

where $F>0$ is the amplitude, ρ is the (angular) frequency, not necessarily positive, and φ_1 the initial phase. All three are assumed to be known constants. We will examine how this oscillation is conveyed from car to car in the line. To start with, consider the first equation (4.12a). Substitute the forcing term into (4.12a) to get

$$\dot{v}_2(t) = F \sin(\rho t - \varphi_1) - \alpha_2 v_2(t-\tau_2). \tag{4.14a}$$

This is a non-homogeneous equation. The associated homogeneous equation is

$$\dot{w}(t) = -\alpha_2 w(t-\tau_2), \tag{4.14b}$$

which has been considered in the last subsection. According to Theorem 6.1 in Hale and Lunel, 1993, the general solution of (4.14a) has the form

$$v_2(t) = w_0(t) + \int_0^t W(t-s)\, g(s)\, \mathrm{d}s,$$

where $w_0(t)$ is the solution of the corresponding homogeneous equation with a given initial condition, $W(t)$ is the fundamental solution of the homogeneous equation, and $g(s)$ is the forcing term. We have seen in the last subsection that the solution of the homogeneous equation can be expressed as linear combinations of (infinite number of) exponential solutions. So we assume

$$W(t) = \Sigma_i\, c_i e^{\lambda_i t}$$

where λ_i are roots of the characteristic equation of the homogeneous equation, and c_i are constants. Substituting this solution and the forcing term into the general solution gives

$$\begin{aligned} v_2(t) &= w_0(t) + \int_0^t \Sigma_i\, c_i e^{\lambda_i (t-s)} F \sin(\rho s - \varphi_1)\, \mathrm{d}s \\ &= w_0(t) + \Sigma_i\, c_i e^{\lambda_i t} \int_0^t F e^{-\lambda_i s} \sin(\rho s - \varphi_1)\, \mathrm{d}s. \end{aligned}$$

The integration can be found to be

$$\int_0^t F e^{-\lambda_i s} \sin(\rho s - \varphi_1)\, \mathrm{d}s = e^{-\lambda_i t}(A_i \sin \rho t + B_i \cos \rho t) - B_i,$$

where

$$\begin{aligned} A_i &= \frac{F}{\lambda_i^2 + \rho^2}(-\lambda_i \cos \varphi_1 - \rho \sin \varphi_1) \\ B_i &= \frac{F}{\lambda_i^2 + \rho^2}(-\rho \cos \varphi_1 + \lambda_i \sin \varphi_1). \end{aligned}$$

Thus the general solution becomes

$$v_2(t) = w_0(t) + \Sigma_i c_i e^{\lambda_i t}\left[e^{-\lambda_i t}(A_i \sin \rho t + B_i \cos \rho t) - B_i\right]$$
$$= w_0(t) - \Sigma_i B_i c_i e^{\lambda_i t} + \Sigma_i c_i (A_i \sin \rho t + B_i \cos \rho t).$$

This solution consists of two parts. The first two terms are linear combinations of solutions of the homogeneous equation, which will die out as $t \to \infty$ if $a_2\tau_2 < \pi/2$, as has shown in the last subsection. The third term is not damped with time and so might be considered as the steady state. The steady state is the linear combination of oscillations with no other frequency than that of the forcing term. Thus for the non-homogeneous equation we can seek a solution which is an oscillation with the same frequency as the forcing term, as follows

$$v_2(t) = r \sin(\rho t - \varphi_1) + s \cos(\rho t - \varphi_1),$$

where r and s are constants to be determined. Then,

$$\begin{aligned}
\dot{v}_2(t) &= r\rho \cos(\rho t - \varphi_1) - s\rho \sin(\rho t - \varphi_1),\\
v_2(t-\tau_2) &= r \sin(\rho(t-\tau_2) - \varphi_1) + s \cos(\rho(t-\tau_2) - \varphi_1)\\
&= r \sin(\rho t - \varphi_1) \cos \rho\tau_2 - r \cos(\rho t - \varphi_1) \sin \rho\tau_2\\
&\quad + s \cos(\rho t - \varphi_1) \cos \rho\tau_2 + s \sin(\rho t - \varphi_1) \sin \rho\tau_2\\
&= (r \cos \rho\tau_2 + s \sin \rho\tau_2) \sin(\rho t - \varphi_1)\\
&\quad - (r \sin \rho\tau_2 - s \cos \rho\tau_2) \cos(\rho t - \varphi_1).
\end{aligned}$$

Substituting these to (4.14a) and equating the coefficients of $\cos(\rho t - \varphi_1)$ and $\sin(\rho t - \varphi_1)$ respectively give

$$\begin{aligned}
-\rho s + a_2 r \cos \rho\tau_2 + a_2 s \sin \rho\tau_2 &= F,\\
\rho r - a_2 r \sin \rho\tau_2 + a_2 s \cos \rho\tau_2 &= 0.
\end{aligned}$$

Solve this set of equations to obtain

$$r = \frac{F\, a_2 \cos \rho\tau_2}{(a_2 \sin \rho\tau_2 - \rho)^2 + a_2^2 \cos^2 \rho\tau_2},$$

$$s = \frac{F\,(a_2 \sin \rho\tau_2 - \rho)}{(a_2 \sin \rho\tau_2 - \rho)^2 + a_2^2 \cos^2 \rho\tau_2}.$$

Therefore, the solution to (4.14a) is

$$v_2(t) = F\frac{a_2\cos\ \rho\tau_2\ \sin\ (\rho t-\varphi_1)\ +\ (a_2\sin\ \rho\tau_2-\rho)\ \cos\ (\rho t-\varphi_1)}{(a_2\sin\ \rho\tau_2-\ \rho)^2\ +\ a_2^2\ \cos^2\rho\tau_2}.$$

To gain a clearer physical meaning of the solution, write the above solution as

$$v_2(t) = \sqrt{r^2\ +\ s^2}\ \sin\ (\rho t-\varphi_1-\arctan\ (-\frac{s}{r})) \tag{4.15}$$

$$= K_2F\ \sin\ (\rho t-\varphi_1-\varphi_2),$$

where

$$K_2 = \frac{1}{\sqrt{(a_2\sin\ \rho\tau_2-\ \rho)^2\ +\ a_2^2\ \cos^2\rho\tau_2}}$$

and

$$\varphi_2 = \arctan\ (-\frac{a_2\sin\ \rho\tau_2\ -\ \rho}{a_2\cos\ \rho\tau_2}).$$

This solution has the same period to that of the forcing term but different amplitude and phase. Besides, when $\alpha_2\tau_2 = \pi/2$, we know the corresponding homogeneous equation has a undamped oscillatory solution with a frequency of $\pi/(2\tau_2)$. In solution (4.15), when $\alpha_2\tau_2 = \pi/2$, and when $\rho \to \pi/(2\tau_2)$, that is, when the frequency of the forcing term approaches the generic frequency of the model, the amplitude of the solution tends to infinity and the system will collapse. This phenomenon is similar to undamped, forced oscillations governed typically by a non-homogeneous linear ordinary differential equation.

Now with the same method as above we can proceed to solve (4.12b), making use of the recurrence relation. First, when $n = 3$, the equation is

$$\dot{v}_3(t) = a_2v_2(t-\tau_2) - a_3v_3(t-\tau_3).$$

Substituting (4.15) into this equation yields

$$\dot{v}_3(t) = a_2K_2F\ \sin\ (\rho t-\varphi_1-\varphi_2-\rho\tau_2) - a_3v_3(t-\tau_3).$$

It is easy to see that the solution of this equation is

$$v_3(t) = K_3\alpha_2K_2F \sin(\rho t-\varphi_1-\varphi_2-\rho\tau_2-\varphi_3),$$

where

$$K_3 = \frac{1}{\sqrt{(\alpha_3 \sin \rho\tau_3 - \rho)^2 + \alpha_3^2 \cos^2\rho\tau_3}}$$

and

$$\varphi_3 = \arctan\left(-\frac{\alpha_3 \sin \rho\tau_3 - \rho}{\alpha_3 \cos \rho\tau_3}\right).$$

Thus, for $n = 3, 4, \cdots, N$, the solution can be written as

$$v_n(t) = K_n\alpha_{n-1}\cdots K_3\alpha_2K_2F \sin(\rho t-\varphi_1-\varphi_2-\rho\tau_2-\varphi_3-\cdots-\rho\tau_{n-1}-\varphi_n), \tag{4.16}$$

where

$$K_i = \frac{1}{\sqrt{(\alpha_i \sin \rho\tau_i - \rho)^2 + \alpha_i^2 \cos^2\rho\tau_i}},$$

and

$$\varphi_i = \arctan\left(-\frac{\alpha_i \sin \rho\tau_i - \rho}{\alpha_i \cos \rho\tau_i}\right),$$

for $i = 2, 3, \cdots, N$.

The linear car-following model can be expressed as

$$\ddot{x}_n(t) = \alpha_n(\dot{x}_{n-1}(t-\tau_n) - \dot{x}_n(t-\tau_n)) = \alpha_n v_n(t-\tau_n), \quad n = 2, 3, \ldots, N.$$

From (4.15) and (4.16), we have

$$\ddot{x}_n(t) = \alpha_n K_n \cdots \alpha_2K_2F \sin(\rho t-\varphi_1-\varphi_2-\rho\tau_2-\varphi_3-\cdots-\varphi_n-\rho\tau_n), \quad n = 2, 3, \ldots, N.$$

Let

$$L_i = a_i K_i \quad \text{and} \quad \psi_i = \varphi_i + \rho\tau_i \quad \text{for } i = 2, 3, \cdots, N.$$

Then

$$\ddot{x}_n(t) = L_n \cdots L_2 F \sin(\rho t - \varphi_1 - \psi_2 - \ldots - \psi_n), \quad n = 2, 3, \cdots, N, \tag{4.17}$$

where

$$L_i = \frac{a_i}{\sqrt{(a_i \sin \rho\tau_i - \rho)^2 + a_i^2 \cos^2 \rho\tau_i}}$$

and

$$\tan \psi_i = \tan(\varphi_i + \rho\tau_i) = \frac{\tan \varphi_i + \tan \rho\tau_i}{1 - \tan \varphi_i \tan \rho\tau_i},$$

or

$$\psi_i = \arctan\left(\frac{\rho \cos \rho\tau_i}{a_i - \rho \sin \rho\tau_i}\right)$$

for $i = 2, 3, \cdots, N$.

By comparing the output $\ddot{x}_n(t)$ in (4.17) with the input $\ddot{x}_1(t)$ in (4.13) it can be seen that the oscillation of the first car is conveyed from car to car with a amplitude coefficient of L_n and a speed of ψ_n per car. The stability over cars can be assured if

$$L_i \leq 1, \quad i = 2, 3, \cdots, N.$$

That is

$$\frac{a_i}{\sqrt{(a_i \sin \rho\tau_i - \rho)^2 + a_i^2 \cos^2 \rho\tau_i}} \leq 1,$$

or

$$\rho^2 - 2a_i \rho \sin \rho\tau_i \geq 0, \quad i = 2, 3, \cdots, N.$$

Multiplying both sides by τ_i gives

$$\rho^2\tau_i - 2a_i\tau_i\rho \sin \rho\tau_i \geq 0$$

or

$$\rho^2\tau_i \geq 2a_i\tau_i\rho \sin \rho\tau_i, \quad i = 2, 3, \cdots, N.$$

This will automatically be satisfied if ρ and $\sin \rho\tau_i$ have different signs so that $\rho \sin \rho\tau_i$ is negative. If ρ and $\sin \rho\tau_i$ are both positive or both negative, then we must have

$$a_i\tau_i \leq \frac{\rho\tau_i}{2\sin \rho\tau_i}, \quad i = 2, 3, \cdots, N.$$

This is the condition for stability over cars. However, in order that the general solution of the equations tends to the oscillating solution found as $t \to \infty$, we must have

$$a_i\tau_i < \pi/2, \qquad i = 2, 3, \ldots, N,$$

as well. Otherwise, if $a_i\tau_i > \pi/2$ for any i, then the general solution of the corresponding equation goes to infinity eventually and there is little point in considering stability over cars. Therefore, to ensure stability over cars, it is necessary that both the above conditions are fulfilled. To get an explicit condition for stability over cars, consider the function

$$h(\rho'_i) \equiv \frac{\rho'_i}{2\sin \rho'_i},$$

where $\rho'_i = \rho\tau_i$. Because ρ and $\sin \rho\tau_i$ are both positive or both negative, $h(\rho'_i)$ must be positive, which implies that

$$\rho'_i \in [0, \pi) \cup (2\pi, 3\pi) \cup (4\pi, 5\pi) \cup \cdots \cup (-\pi, 0] \cup (-3\pi, -2\pi) \cup (-5\pi, -4\pi) \cup \cdots.$$

Within this domain, the following properties of the function can be easily verified:

(1) If $\rho'_i \to 0$, then $h(\rho'_i) \to 1/2 < \pi/2$;

(2) $\dfrac{dh}{d\rho'_i} > 0, \qquad 0 < \rho'_i < \pi$;

(3) If $\rho'_i \to \pi - 0$, then $h(\rho'_i) \to \infty > \pi/2$;

(4) $h(2.3138) = 1.5708 \cong \pi/2$;

(5) $h(\rho'_i) > \pi/2$ if $\rho'_i > 2\pi$. This is because if ρ'_i is increased by 2π, then $h(\rho'_i)$ changes to

$$\frac{\rho'_i + 2\pi}{2\sin \rho'_i} = \frac{\rho'_i}{2\sin \rho'_i} + \frac{2\pi}{2\sin \rho'_i} > 1/2 + \pi > \pi/2.$$

These properties mean that

$$h(\rho'_i) < \pi/2 \text{ if } \rho'_i < 2.3138,$$
$$h(\rho'_i) = 1.5708 \cong \pi/2 \text{ if } \rho'_i = 2.3138,$$

and

$$h(\rho'_i) > \pi/2 \text{ if } \rho'_i > 2.3138.$$

These properties have been listed for $\rho'_i > 0$. With obvious modifications, the properties and the conclusions still hold if $\rho'_i < 0$ because the function is symmetric about $\rho'_i = 0$. Thus the explicit condition for the stability over cars can be stated as follows.

$$\alpha_i \tau_i \leq \frac{\rho\tau_i}{2\sin \rho\tau_i}, \qquad \text{if } |\rho\tau_i| < 2.3138;$$
$$\alpha_i \tau_i < \pi/2, \qquad \text{if } |\rho\tau_i| \geq 2.3138,$$

for $i = 2, 3, \cdots, N$.

A sufficient condition which is simpler and stronger can also be obtained by observing that

$$h(\rho'_i) \geq 1/2.$$

So the two conditions for the stability over cars will all be satisfied if

$$\alpha_i \tau_i \leq 1/2, \qquad i = 2, 3, \cdots, N.$$

This sufficient condition for the stability over cars is similar to that in Chandler *et al.* (1958), with $\alpha\tau \leq 1/2$ being generalized here to $\alpha_i \tau_i \leq 1/2$, $i = 2, 3, \cdots, N$. However, in Chandler *et al.* (1958) different values of $\rho\tau_i$ was not considered explicitly; only the case where $\rho \to 0$ was considered, based on which only the sufficient condition was obtained.

The condition for stability over cars for the non-autonomous model can be extended to consider the stability over cars for the autonomous model. From the last subsection we know that oscillations, both damped and undamped, are typical solutions in the autonomous model. The oscillation of one car can be considered as an input or a forcing term to the next car. Since the frequency of the oscillations may be of any value, the condition $a_i\tau_i \leq 1/2,\ i = 2, 3, \cdots, N$ can be used as a sufficient condition for the stability over cars for the linear autonomous model. This sufficient condition for the stability over cars is stronger than the condition for the stability (with time) of the equilibrium in the linear autonomous model ($a_i\tau_i \leq \pi/2,\ i = 2, 3, \cdots, N$). Therefore, it can guarantee the stability with time as well.

Having derived the conditions for the stability over cars, it should be commented that the instability over cars also depends on the number of cars concerned: it is the oscillation in the motion of the last car that has the biggest amplitude when the system is unstable over cars. If there are only two or three cars in the line, then the system may still run steadily even if the criterion for stability over cars is violated.

4.2.3. The nonlinear autonomous model

Setting $\ddot{x}_1(t) = 0$ in (4.3) we get the nonlinear autonomous model

$$\dot{v}_2(t) = -\beta_2(t)\, v_2(t-\tau_2), \tag{4.18a}$$

$$\dot{v}_n(t) = \beta_{n-1}(t)\, v_{n-1}(t-\tau_{n-1}) - \beta_n(t)\, v_n(t-\tau_n), \quad n = 3, 4, \ldots, N, \tag{4.18b}$$

$$\dot{y}_n(t) = v_n(t), \quad n = 2, 3, \ldots, N, \tag{4.18c}$$

where

$$\beta_n(t) = \frac{a_n(\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau_n) + b)^l}, \quad n = 2, 3, \ldots, N.$$

In this model, there is an equilibrium where

$$v_n^e = 0, \quad y_n^e = a_n, \quad n = 2, \ldots, N, \tag{4.19}$$

where a_n is some constant, the superscripts "e" denote equilibrium. This is a continuum of equilibria because it exists for any value of $a_n \geq 0$; the limiting relative spacing depends on the initial conditions. We consider the stability of the equilibria in this subsection.

Stability analysis is difficult for nonlinear systems, especially for delay-differential equations. There are hardly any stability analyses of the nonlinear car-following model. An ordinary way to investigate the stability of an equilibrium in a nonlinear system is to consider the local stability by linearizing the model at the equilibrium (Kuang, 1993). This method will be used here.

First of all, we can write (4.18) as

$$\begin{aligned}
\dot{v}_2(t) &= -\ddot{x}_2(t), \\
\dot{v}_n(t) &= \ddot{x}_{n-1}(t) - \ddot{x}_n(t), \quad n = 3, 4, \ldots, N, \\
\dot{y}_n(t) &= v_n(t), \quad n = 2, 3, \ldots, N,
\end{aligned}$$

where

$$\begin{aligned}
\ddot{x}_n(t) &= f_n(v_2(t), \cdots, v_n(t), v_n(t-\tau_n), y_n(t-\tau_n)) \\
&= \beta_n(t)\, v_n(t-\tau_n)
\end{aligned}$$

Consider a small disturbance, $\delta v_n(t)$ and $\delta y_n(t)$, superimposed on the equilibrium:

$$\begin{aligned}
v_n(t) &= v_n^e + \delta v_n(t), \quad n = 2, 3, \ldots, N, \\
y_n(t) &= y_n^e + \delta y_n(t), \quad n = 2, 3, \ldots, N.
\end{aligned}$$

To see how these small disturbances evolve, substitute these into (4.18) and linearize round the equilibrium point by taking the first order terms of the Taylor expansion of the equations. Then

$$\delta \dot{v}_2(t) = -\Delta f_2^e,$$

$$\delta \dot{v}_n(t) = \Delta f^e_{n-1} - \Delta f^e_n, \quad n = 3, 4, \ldots, N,$$
$$\delta \dot{y}_n(t) = \delta v_n(t), \qquad n = 2, 3, \ldots, N,$$

where

$$\Delta f^e_n = \left[\delta v_2(t) \frac{\partial}{\partial v_2} + \ldots + \delta v_n(t) \frac{\partial}{\partial v_n} + \delta v_n(t-\tau_n) \frac{\partial}{\partial v_n(t-\tau_n)} \right.$$
$$\left. + \delta y_n(t-\tau_n) \frac{\partial}{\partial y_n(t-\tau_n)} \right] f^e_n$$

and

$$f^e_n = f_n(v^e_2(t), \cdots, v^e_n(t), v^e_n(t-\tau_n), y^e_n(t-\tau_n)).$$

The derivatives can be found to be

$$\frac{\partial f_i}{\partial v_j} = -\frac{m\, a_n(\dot{x}_1(t) - v_2(t) - \cdots - v_i(t))^{m-1}}{(y_i(t-\tau_i) + b)^l} v_i(t-\tau_i),$$
$$j = 2, 3, \ldots, i; \quad i = 2, 3, \ldots, n,$$

$$\frac{\partial f_n}{\partial v_n(t-\tau_n)} = \frac{a_n(\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau_n) + b)^l}, \quad n = 2, 3, \ldots, N,$$

$$\frac{\partial f_n}{\partial y_n(t-\tau_n)} = -l \frac{a_n(\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau_n) + b)^{l+1}} v_n(t-\tau_n),$$
$$n = 2, 3, \ldots, N.$$

Evaluating the derivatives at the equilibrium (4.19), we have

$$\delta \dot{v}_2(t) = -\frac{a_2\, \dot{x}_1^m}{(y^e_2 + b)^l} \delta v_2(t-\tau_2),$$
$$\delta \dot{v}_n(t) = \frac{a_{n-1}\, \dot{x}_1^m}{(y^e_{n-1} + b)^l} \delta v_{n-1}(t-\tau_{n-1}) - \frac{a_n\, \dot{x}_1^m}{(y^e_n + b)^l} \delta v_n(t-\tau_n),$$
$$n = 3, 4, \ldots, N$$

$$\delta \dot{y}_n(t) = \delta v_n(t) \qquad n = 2, 3, \ldots, N$$

Denote the value of sensitivity of nth driver at the equilibrium point by

$$\beta_n^e = \frac{a_n \dot{x}_1^m}{(y_n^e + b)^l}.$$

Then

$$\begin{aligned}
\delta \dot{v}_2(t) &= -\beta_2^e \delta v_2(t-\tau_2), \\
\delta \dot{v}_n(t) &= \beta_{n-1}^e \delta v_{n-1}(t-\tau_{n-1}) - \beta_n^e \delta v_n(t-\tau_n), \quad n = 3, 4, \ldots, N \\
\delta \dot{y}_n(t) &= \delta v_n(t), \qquad n = 2, 3, \ldots, N.
\end{aligned} \tag{4.20}$$

This set of equations has the same form as that of the linear autonomous model (4.4) and so has an equilibrium

$$\delta v_n(t) \equiv 0, \ \delta y_n(t) \equiv \text{some constant}, \quad n = 2, 3, \ldots, N \tag{4.21}$$

The local stability of the equilibrium in the nonlinear model follows from that of its linearized model. From the analysis in 4.2.1, the local stability of the equilibrium can be summarized as follows:

(1) When $\max_n (\beta_n^e \tau_n) > \pi/2$, disturbances are amplified oscillations and the equilibrium is nonstable.

(2) When $\max_n (\beta_n^e \tau_n) = \pi/2$, disturbances are conserved and the equilibrium is neutrally stable.

(3) When $1/e < \max_n (\beta_n^e \tau_n) < \pi/2$, disturbances are damped with oscillations and the equilibrium is stable.

(4) When $\max_n (\beta_n^e \tau_n) \leq 1/e$ disturbances are damped with oscillations of frequencies higher than π only and the equilibrium is stable.

While $\max_n \beta_n^e \tau_n$ can determine the stability of the equilibrium, it is not a single constant since β_n^e depends on y_n^e, which in turn, depends on initial conditions (Note that for an autonomous model, the speed of the first car, $\dot{x}_1$, is a known constant). Therefore, the values of parameters for an equilibrium to become unstable are different for different initial conditions.

4.3. NUMERICAL ANALYSIS

4.3.1. The algorithm

The essential part of numerical analysis is the integration of the car-following model. There is no standard algorithm to integrate a delay-differential equation. However, an algorithm for integrating ordinary differential equations can be modified to solve delay-differential equations.

There are three types of commonly used algorithms for integrating ordinary differential equations, the *Runge-Kutta methods* which are one step methods, the *multistep methods*, and the *predictor corrector methods*. They all have different advantages and disadvantages; no one type of method performs uniformly better than another type for all purposes. The multistep methods and the predictor corrector methods are more efficient and give error estimates. However, they are not self-starting; an independent method is needed to obtain starting values, and to obtain necessary values for predictor corrector methods when the step size is doubled or halved. Besides, the multistep methods are more complicated to program and may be subject to numerical instability. The Runge-Kutta methods have an important advantage that they are self-starting and consequently are straightforward to program. In addition, they are numerically stable. They provide good accuracy and occupy relatively a small amount of computer storage.

To solve the car-following model, a standard algorithm for ordinary differential equations has to be modified to deal with the delay term. So, a simpler method is desired. Therefore the fourth order Runge-Kutta method will be considered, though there are higher order methods. Farmer (1982) modified the standard Runge-Kutta algorithm to solve a delay-differential equation. The modification used here is the same as Farmer's (Farmer,1982), but the algorithm is adapted so as to integrate the system of car-following equations. The Runge-Kutta method

and its modifications can be found in Appendix A. The algorithm is coded in FORTRAN and the program is listed in Appendix B.1. In order to test the program, it is used to integrate the delay-differential equation modelling blood production due to Makey and Glass (Farmer, 1982)

$$\dot{X} = \frac{0.2\ X_\tau}{1 + X_\tau^{10}} - 0.1\ X,$$

where X is the concentration of blood, τ is the delay, and $X_\tau \equiv X(t-\tau)$. There is a strange attractor in this model when $\tau = 23$ (Farmer, 1982). The model is integrated for the strange attractor and the solution plotted in Figure 4.4. The attractor has the same pattern as that in Farmer (1982). Note that on a strange attractor two solutions follow the same trajectory only if they have *exactly* the same initial conditions.

The program is also used to integrate the linear non-autonomous 2-car model (4.14a) for which the analytic solution has been found in 4.2.2. The numerical and analytical solutions are compared in Figure 4.5, where the circles are the analytical solution and the solid line is the numerical solution. Good agreement can be seen clearly in this diagram.

In the next two subsections the nonlinear car-following model (4.2) is investigated numerically. The model is copied here:

$$\begin{aligned}
\dot{v}_2(t) &= \ddot{x}_1(t) - \beta_2(t)\ v_2(t-\tau), \\
\dot{v}_n(t) &= \beta_{n-1}(t)\ v_{n-1}(t-\tau) - \beta_n(t)\ v_n(t-\tau), \quad n = 3, 4, \ldots, N, \\
\dot{y}_n(t) &= v_n(t), \quad n = 2, 3, \ldots, N,
\end{aligned}$$

where

$$\beta_n(t) = \frac{a\ (\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau) + b)^l}, \quad n = 2, 3, \ldots, N.$$

The forcing term considered here is (4.13), used in the theoretical analysis of the linear non-autonomous model. The initial phase is not significant, so we take

$$\ddot{x}_1(t) = F \sin \rho t.$$

Then speed of the first car is

$$\dot{x}_1(t) = \bar{u} - \frac{F}{\rho}\cos \rho t,$$

where $\bar{u}$ is the average speed of the first car. If $F = 0$, then $\ddot{x}_1(t) \equiv 0$, $\dot{x}_1(t) \equiv \bar{u}$, and the model is autonomous. If $F > 0$, then the model is non-autonomous. In the numerical study, the car-following equations are integrated until a steady state is approached. The solutions can then be visualized by plotting the time series and the phase portrait projections.

4.3.2. The dynamic behaviour of the nonlinear model

In this analysis, different initial conditions and values of parameters are tested to identify possible attractors and to find out how they change with values of parameters. The autonomous and the non-autonomous models are tested respectively. The results of calculations can be described as follows.

The autonomous model

When there is no forcing term, two types of attractor are found: a continuum of equilibria where the relative speed is zero and the relative spacing constant, and periodic attractors. The values of the parameters determine which attractor the solutions approach. It seems that the product of the sensitivity at the equilibrium and the reaction time, or

$$\beta_n^e \tau = \frac{\alpha \bar{u}^m}{(y_n^e + b)^l}\tau$$

functions as a whole to determine qualitative properties of the behaviour of the model. This is consistent with the theoretical analysis in section 4.2.3. Of course $\beta_n^e \tau$ is unknown before each run of the system because y_n^e is unknown. However, y_n^e depends on initial conditions which can be controlled in the experiments. For given initial conditions, if $\beta_n^e \tau$ is small, as a result of smaller

α, τ, m, and $\bar{u}$, or bigger l, and b, the solutions always approach stable equilibria. As $\beta_n^e\tau$ is increased gradually (through increasing of α, τ, m, and $\bar{u}$, or decreasing of l, and b), the equilibrium becomes unstable and a stable periodic orbit emerges. Figure 4.6 shows the convergence of one such periodic solution in the autonomous model.

The frequency of the periodic solutions in the autonomous model is examined numerically. It can be seen in Figure 4.6 the frequency of the periodic solution is $\pi/(2\tau)$. Note that a frequency of $\pi/(2\tau)$ is equivalent to a period of 4τ seconds; the unit of time is second in this figure. In fact, it has been found from a lot of calculations that the frequencies of oscillations are all round $\pi/(2\tau)$, independent of value of parameters and the initial conditions. Only amplitudes depend on these factors.

The ranges of values of parameters and initial conditions for the model to have a periodic attractor are rather narrow in most cases experimented. Increasing sensitivity and/or reaction time often causes the system to break down with the speed of the last car being negative, even before a periodic solution occurs. The break down is caused by the instability over cars: when disturbances are amplified along cars it is always the solution of the last car that has the largest amplitude and so collapses first.

Figures 4.7a–4.7b are produced to show the amplification of disturbances along cars. In these two figures, the solid line is the relative speed of the second car to the first car, the dashed line is the relative speed of the sixth car to the fifth car, and the dotted line is the relative speed of the tenth car to the ninth car. It can be seen that although disturbances are damped with time for each car, they are amplified along cars. In addition, the amplification is more serious when α is larger (Figure 4.7b). In the theoretical analysis in the last section, we showed that for the linear autonomous model the stability over cars may need a stronger condition than the stability with time does. Here in Figure 4.7 we have a nonlinear autonomous model that is stable with time but not stable over cars.

The breaking down can be removed by taking away the last car, but it will be bound to occur at further increases of sensitivity or reaction time. It may also be "introduced" by increasing the number of cars in the line, with or without increasing of sensitivity or reaction time. The reason for this is obvious. The

stability over cars depends on the sensitivity and the reaction time; the occurrence of instability over cars depends on the number of cars in the line as well. If the number of cars is small, the amplitude of the oscillation of the last car may not cause the system to break down even though it is larger than that of the car in front.

The non-autonomous model

When the forcing term is introduced, the model behaves in a very similar way to that of autonomous model. If the reaction time and/or the sensitivity is small enough, then the stable equilibria in the autonomous model are replaced by stable periodic orbits with the same period as that of the forcing term. The amplitudes of the oscillations depend on the initial conditions, the reaction time, the sensitivity, and the position of the car in the line. It seems that for small values of the reaction time or the sensitivity the disturbances are damped with time, leaving the forcing oscillations as the steady states. This feature is similar to that of the linear non-autonomous car-following model considered in 4.2.2.

When the reaction time or the sensitivity is larger, we know that at some point the autonomous model has a periodic solution with a frequency of about $\pi/(2\tau)$. If a forcing term is added in this situation then the model has a more complicated behavior. Most of the steady-state solutions, however, are still periodic, although the frequency of the solutions may not be the same as that of the forcing term.

Figures 4.8 to 4.10 show three seemingly complicated solutions of the non-autonomous model. However, in each case, examination of the power spectrum and a longer plot showed that the solution is periodic. Figure 4.11 shows another solution found in the non-autonomous model. The relative speed and spacing of the third car to the fourth car are plotted here. This solution, though it looks complicated, is not chaotic: it appears to be quasi-periodic, or possibly periodic with a very long period. As mentioned in Chapter 3, in practice it is impossible to be sure that a solution is quasi-periodic rather than periodic.

The seemingly complicated behaviour in the non-autonomous model may occur only when the corresponding autonomous model has a periodical solution and only for limited ranges of values of parameters and initial conditions. Again, increasing the sensitivity or reaction time often causes instability over cars in the

system: the system collapses in the same way as that for the autonomous model. Figures 4.12a-4.12b show the convey of disturbances along cars in a nonlinear non-autonomous model. In these two figures, the solid line is the relative speed of the second car to the first car, the dashed line the relative speed of the third car to the second car, and the dotted line the relative speed of the fourth car to the third car. In Figure 4.12a, $\alpha = 110$ and the oscillation of the forcing term is damped along cars. In Figure 4.12b, $\alpha = 130$ and the oscillation is amplified.

Summary of numerical calculations

Many calculations have been carried out and have shown that the phenomena outlined above about the nonlinear model, both autonomous and non-autonomous, are typical in the car-following model. No other feature has emerged.

Table 4.1 shows the results of experiments on the parameters α, τ, m, and l with a 2-car model. Each parameter was tested separately with all the other parameters fixed. The values of parameters tested ranging from the value at which driver's reactions are very weak and the system is mild, to the value at which the system is excited and then collapses afterwards. The table can give a general view of the dynamics of the car-following model.

In this table, the upper half shows the values of parameters experimented and the lower half the values of parameters (ranges or points) for different behaviour in the car-following model. The average speed of the first car $\bar{u}$ is chosen to be one of a typical value of driving speed; the minimum spacing b is determined from $\bar{u}$ according to driving regulations (The Department of Transport, 1988). The tests are made for both absence and presence of the forcing term. The results of the Table can be summarized as follows. When the value of parameter is small for α, τ, m, or large for l, the steady-state solutions are equilibria for the autonomous model and periodic with the same frequency as that of the forcing term for the non-autonomous model. As parameters increase (or decrease in case of l) such that $\beta^e\tau$ is about $\pi/2$, the solutions become periodic for autonomous model. (This stability property of the equilibria in the autonomous model is similar to the local stability property found in the theoretical analysis, which implies that the linearization gives a good approximation of the nonlinear model).

When the forcing term is added in this situation, the solutions are still periodic in most cases although they may be more complicated. The range of parameters for autonomous models to have periodic solutions and for the non-autonomous models to have more complicated solutions are very limited; the system soon collapses because of the instability over cars if parameters continue to increase (or to decrease for l).

4.4. SUMMARY AND COMMENTS

In this chapter, the car-following model has been investigated both theoretically and numerically. The stability (with time) and the stability over cars in the linear model, both autonomous and non-autonomous, are examined theoretically. The local stability of the nonlinear autonomous model is considered by linearization. The analyses are extended to allow the sensitivity coefficient α and the reaction time τ to be different for different drivers. Criteria for the stabilities are established. The stability conditions for the linear model are more general than those in the literature. Oscillatory solutions with higher frequencies than $\pi/2$ are considered explicitly in stability analysis of the linear autonomous model; different values of the frequency in the forcing term are considered and the condition for stability over cars found is sufficient as well as necessary. The nonlinear model is investigated numerically. The results of the nonlinear autonomous model agree well with the theoretical analysis, indicating that the linearization gives a good approximation to the nonlinear model.

Typical steady states in the car-following model are equilibria and periodic solutions in the autonomous model, and periodic solutions in the non-autonomous model. It is shown that the dynamic properties of the model depend on the sensitivity and the reaction time. This is not unexpected and has its theoretical reason and practical implications. From the car-following equations it can be seen that drivers try to match their speeds with that of the cars in front with a reaction time: they accelerate when they are slower and decelerate when they are faster. When the sensitivity or reaction time is larger, the differences of speeds tend to be over compensated and oscillations emerge. If it is even larger, oscillations can never die down and can even be amplified, either with time or along cars. A moderate and quick reaction to changes of motion of the proceeding car corresponds to a normal and safe driving, while a reaction that is too strong and too slow is a typical behaviour of drinking driving.

Theoretically, the car-following model is complicated because it has a delay time built in and it is nonlinear. The dynamic behaviour of the model, however, is not as complicated as one would expect. As mentioned in Chapter 2, Disbro and Frame (1990) calculated (positive) Liapunov exponents of solutions of the car-following model without showing a chaotic attractor. This is not sufficient to prove that the solutions are chaotic. Here, many numerical calculations have been made in an attempt to identify chaos in the model. In none of the investigations carried out has any evidence of chaos, such as a chaotic attractor, or bifurcation sequences which may lead to chaos, been found. This outcome agrees with that of Kirby and Smith (1991). Of course, no general conclusion can be reached from numerical experiments which can be comprehensive but are necessarily limited . However, at least chaos does not appear to occur typically in the model. Although the periodic solutions in the autonomous model may lead to complicated behaviour when a forcing term is added, they may occur only for small ranges of values of parameters and initial conditions and only when the sensitivity or reaction time is large enough. In most cases, there is not much room for values of parameters to change before the occurrence of collapses.

The car-following model may not be a satisfactory model to describe a car-following process. The real car-following process is very complicated. It involves human behavior and psychological effects which are difficult to model. The car-following model has some unreasonable implications. First, the following driver reacts only to differences in the speeds. If the relative speed is zero then they pay no attention to spacing, however large or small it is! Secondly, drivers' reaction in accelerations is the same as that in decelerations. In reality, one would expect that the response of the following car may depend on whether it is accelerating or decelerating. There have been only limited numbers of empirical studies of the car-following model (Chandler *et al.*, 1958, Gazis *et al.*, 1961). Gazis *et al.* (1961) concluded from some experimental studies that a nonlinear car-following model is necessary to account for the observed traffic flow behaviour, but no particular model had been found to be better than all others. Therefore, empirical studies are needed to find out the values of parameters in the car-following model or even better forms of a car-following model so that any method for the automatic control of driving based on the car–following model can have a sound basis.

Table 4.1 Summary of the behaviour of the car-following model

		parameters tested			
		α	τ	m	l
other parameters	α	2.0^s	3.0	1.0	1.0
	τ	1.0	0.5^s	1.0	1.0
	m	1.5	1.5	1.0^s	1.5
	l	1.5	1.5	1.5	3.0^s
	$\bar{u}$	22	22	22	22
	b	53	53	53	53
	F	1	1	1	1
	ρ	$2\pi/10$	$2\pi/10$	$2\pi/10$	$2\pi/10$
		results			
autonomous model	equl	2.00-5.70	0.50-1.85	1.00-2.06	3.00-1.06
	c_{ep}	5.70-5.75	1.85-1.86	2.06-2.08	1.06-1.05
	$\beta^e\tau$	1.552	1.544	1.537	1.554
	pera	5.75-8.30	1.86-2.15	2.08-2.12	1.05-0.95
	c_{pc}	8.35	2.16	2.14	0.96
non-autonomous model	pern	2.00-5.85	0.50-1.96	1.00-2.06	3.00-1.06
	c_{pq}	5.85-5.90	1.96-1.98	2.06-2.08	1.06-1.05
	qup.	5.90-7.80	1.98-2.16	2.08-2.12	1.05-0.97
	c_{qc}	7.85	2.18	2.14	0.98

Notation:

The superscript ^{s} means the minimum value tested, beyond which the response of the following car is very weak.

equl — equilibria in the autonomous model.

c_{ep} — changing from an equilibrium to periodic orbits in an autonomous model.

$\beta^e\tau$ — the product of sensitivity and the reaction time at the equilibria in an autonomous model.

pera — periodic orbits in the autonomous model.

pern — periodic orbits in the non-autonomous model, whose frequency are the same as that of the forcing term

c_{pq} — changing from periodic solutions with the same frequency as that of the forcing term to a more complicated periodic solution whose frequency may not be the same as that of the forcing term in a non-autonomous model.

c_{pc} — breaking down in an autonomous system.

qup. — more complicated, but still periodic, solutions.

c_{qc} — breaking down in a non-autonomous model.

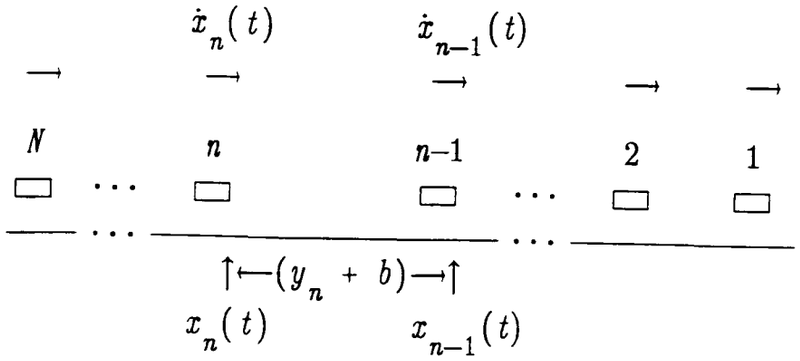


Figure 4.1 The car-following model

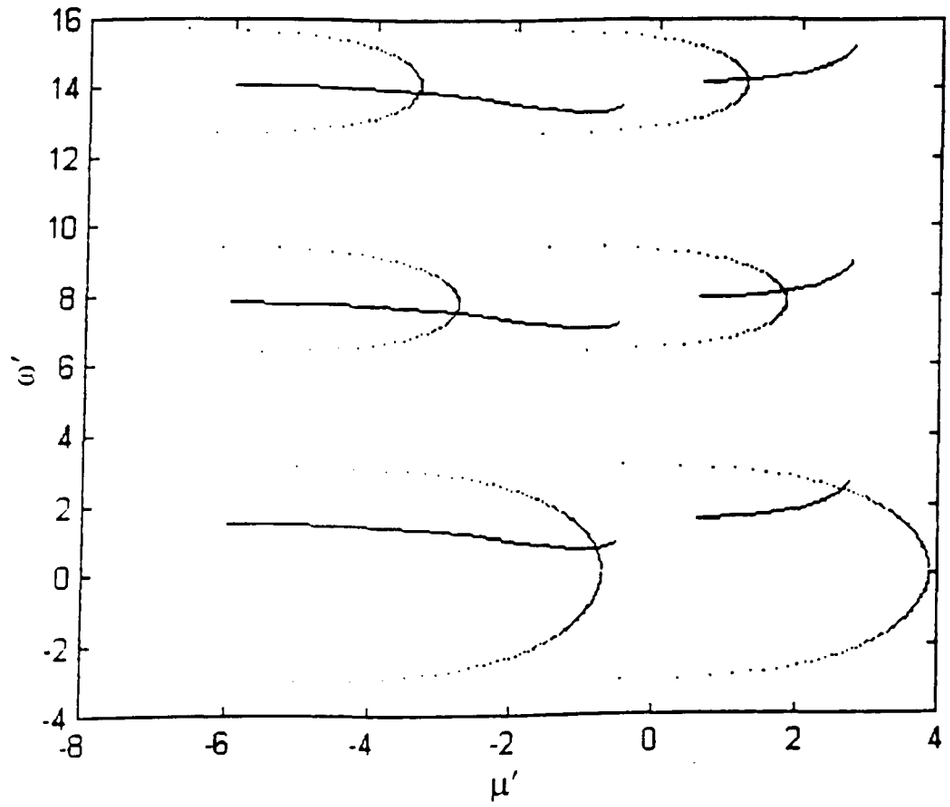
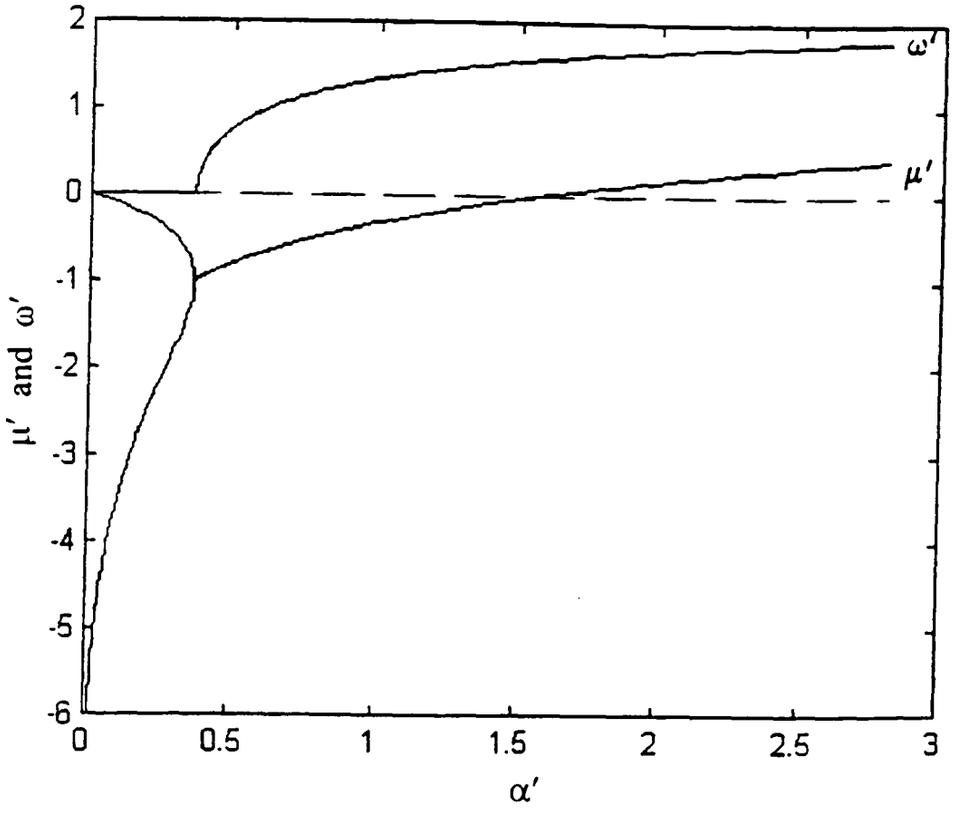
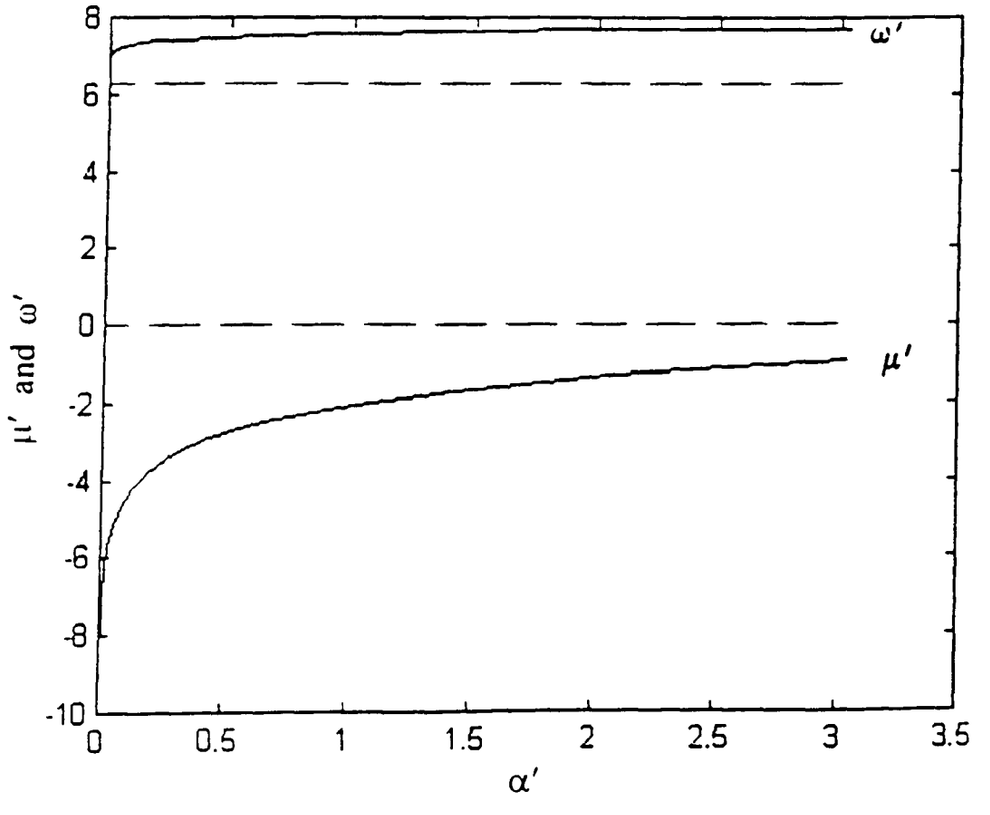


Figure 4.2 Solution of the characteristic equation, with α' equal to 0.5 (left) and 50 (right), and for ω' within $(-\pi, \pi)$, $(2\pi, 3\pi)$, and $(4\pi, 5\pi)$.

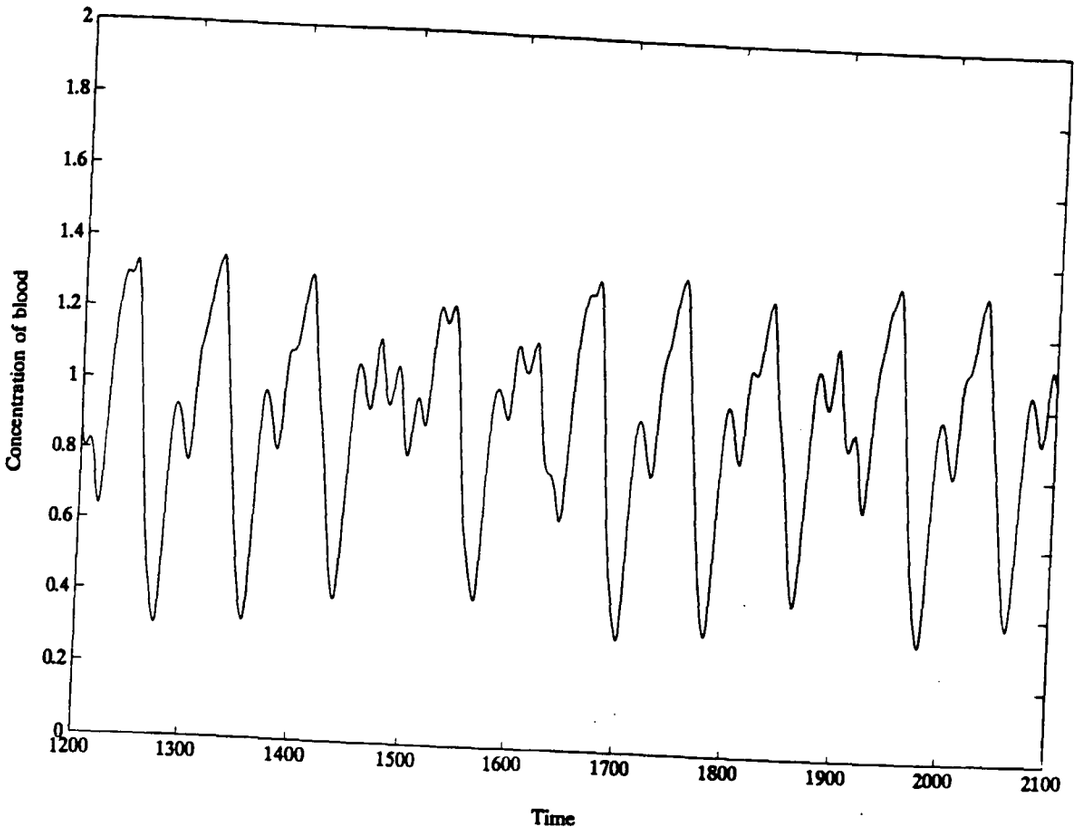


(a)

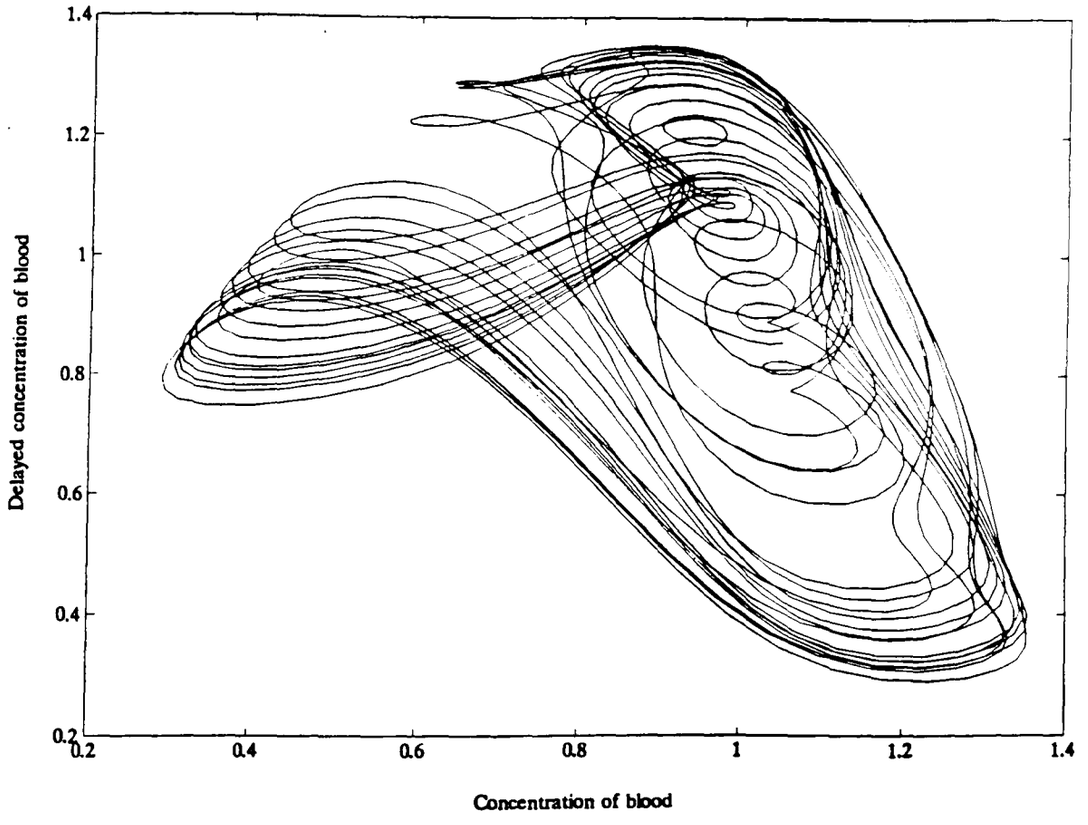


(b)

Figure 4.3 The characteristic equation, the effect of α'. (a) $\omega' \in [0, \pi)$, (b) $\omega' \in (2\pi, 3\pi)$.



(a)



(b)

Figure 4.4 Chaotic attractor of the Makey-Glass equation with $\tau = 23$. (a) Concentration of blood; (b) delayed concentration of blood against the concentration of blood.

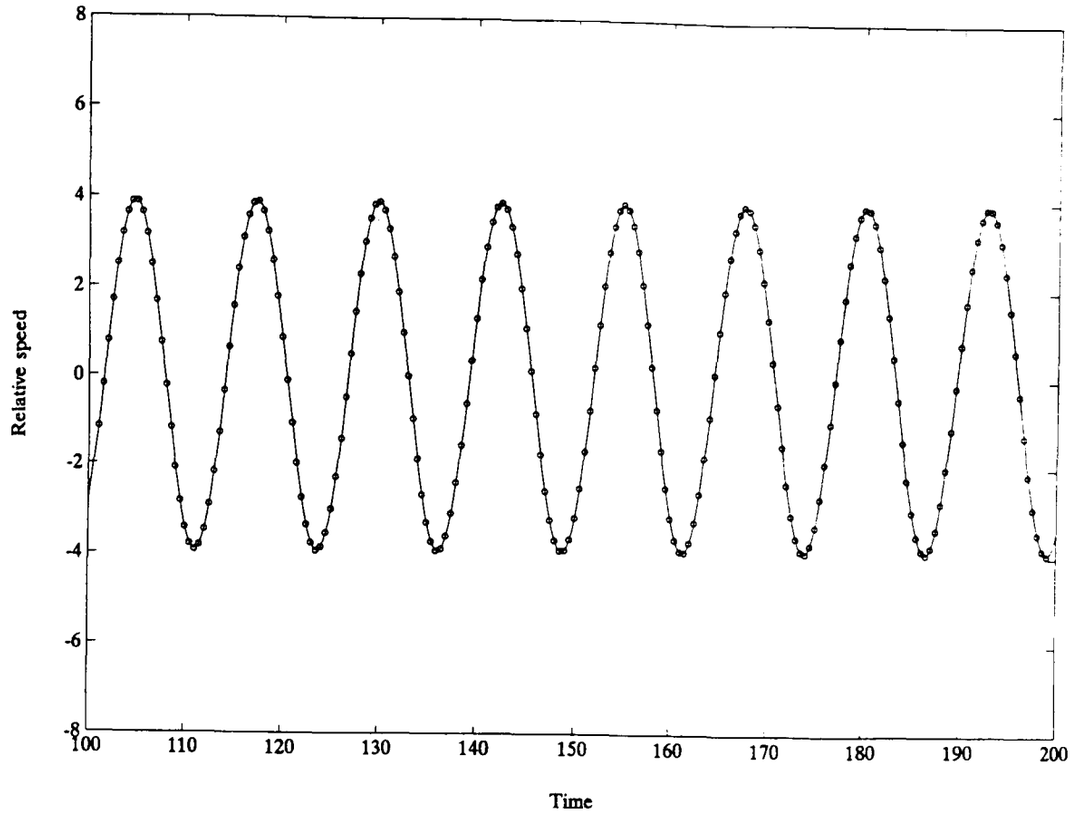


Figure 4.5 Solution of the linear non-autonomous 2-car model (4.14a), with $\alpha_2 = 0.5$, $\tau_2 = 1.0$, $F = 2$, $\rho = 0.5$, $\varphi_1 = 0$.

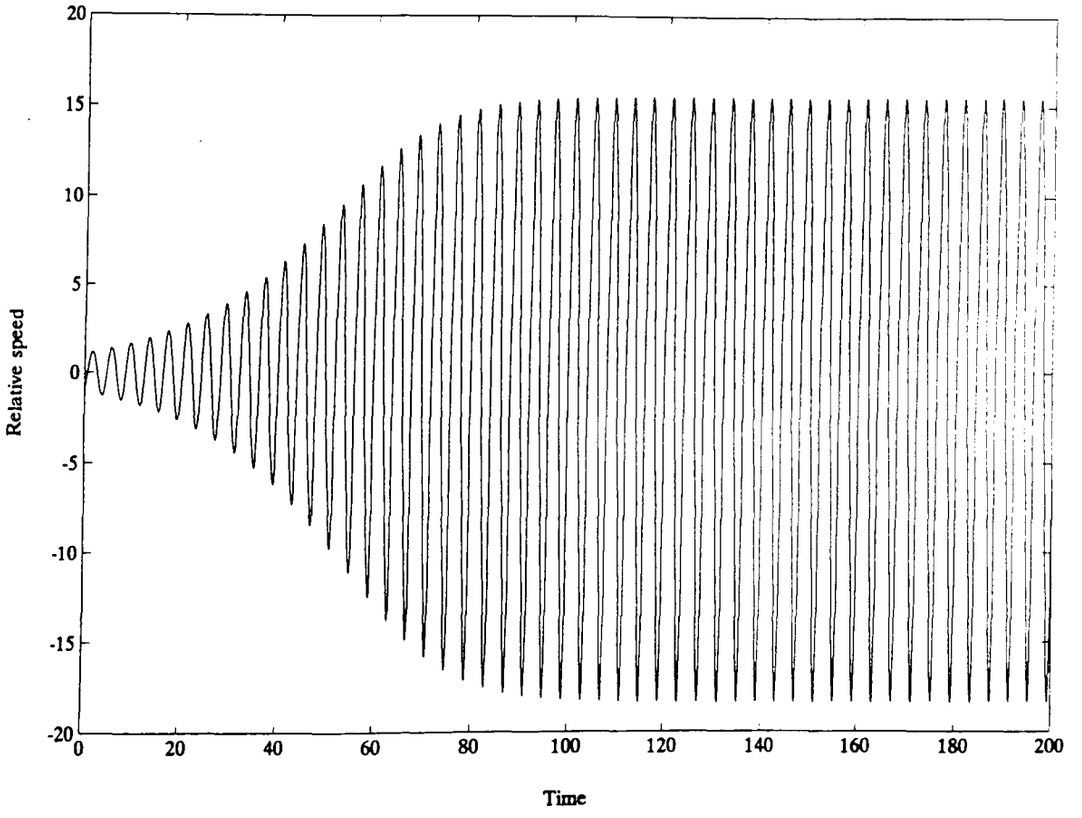
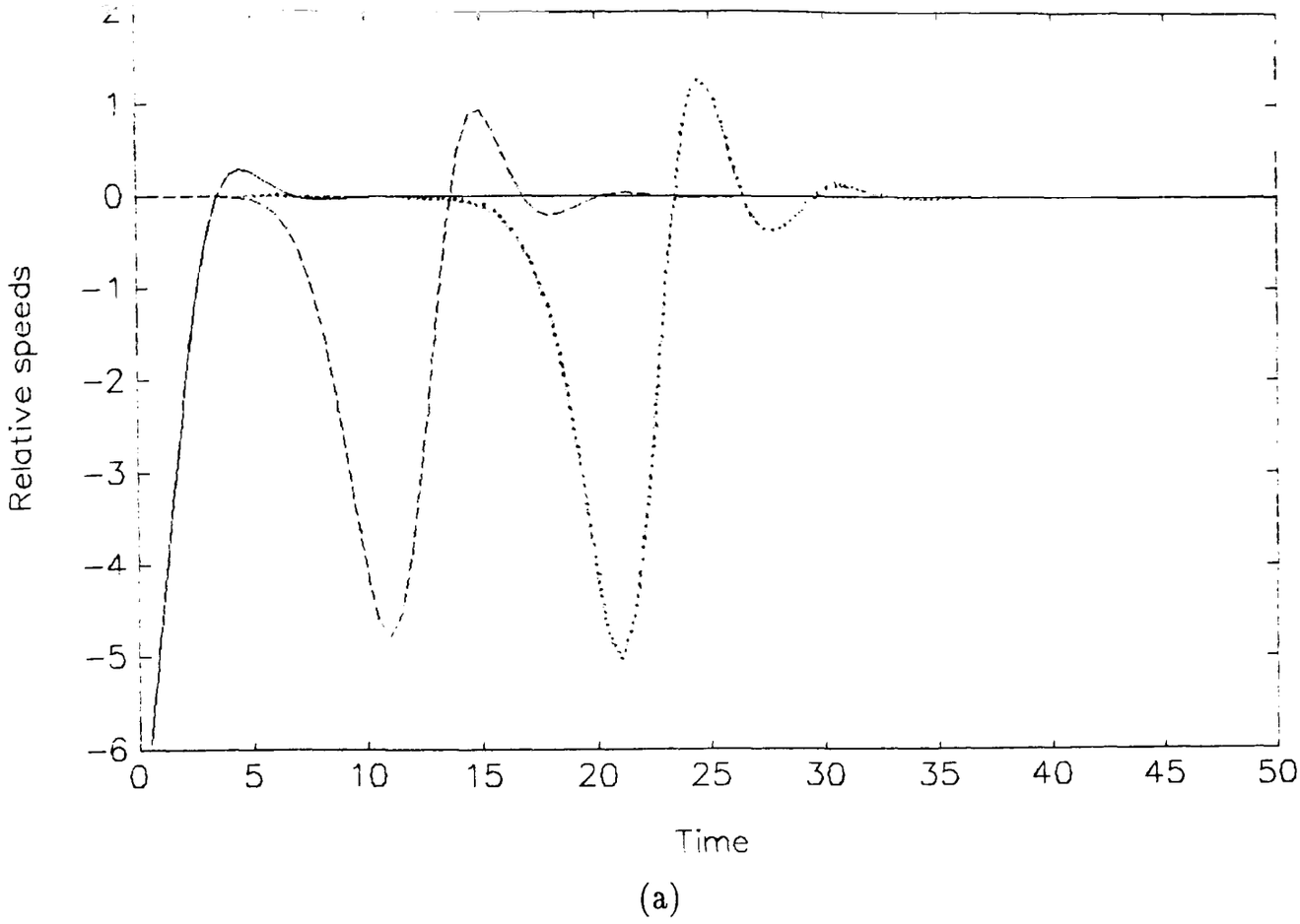
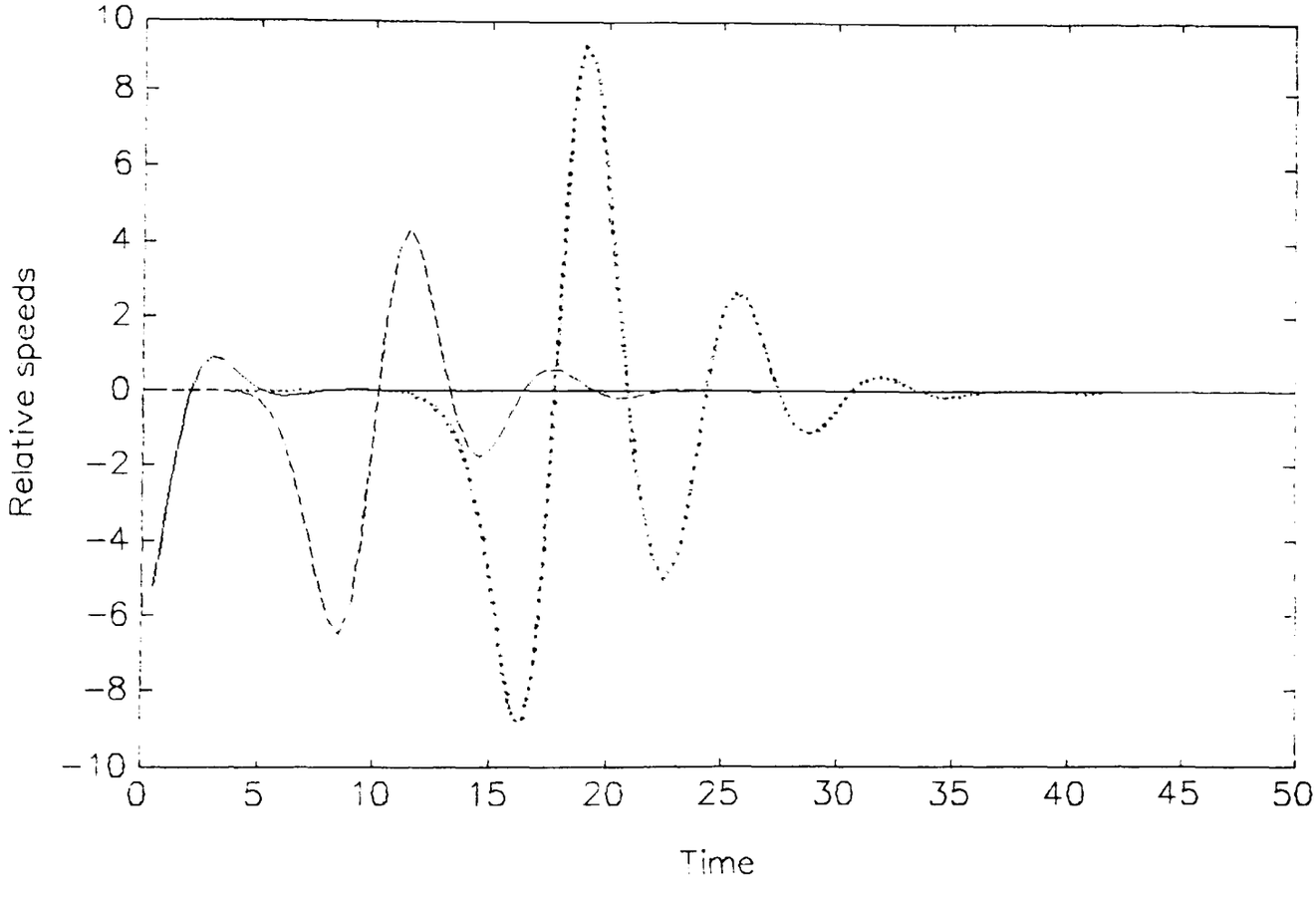


Figure 4.6 Convergence to a periodic attractor in the nonlinear autonomous car-following model, with $N = 2$, $l = 2$, $m = 1$, $\alpha = 130$, $\tau = 1$, $b = 30$, $\bar{u} = 20$.

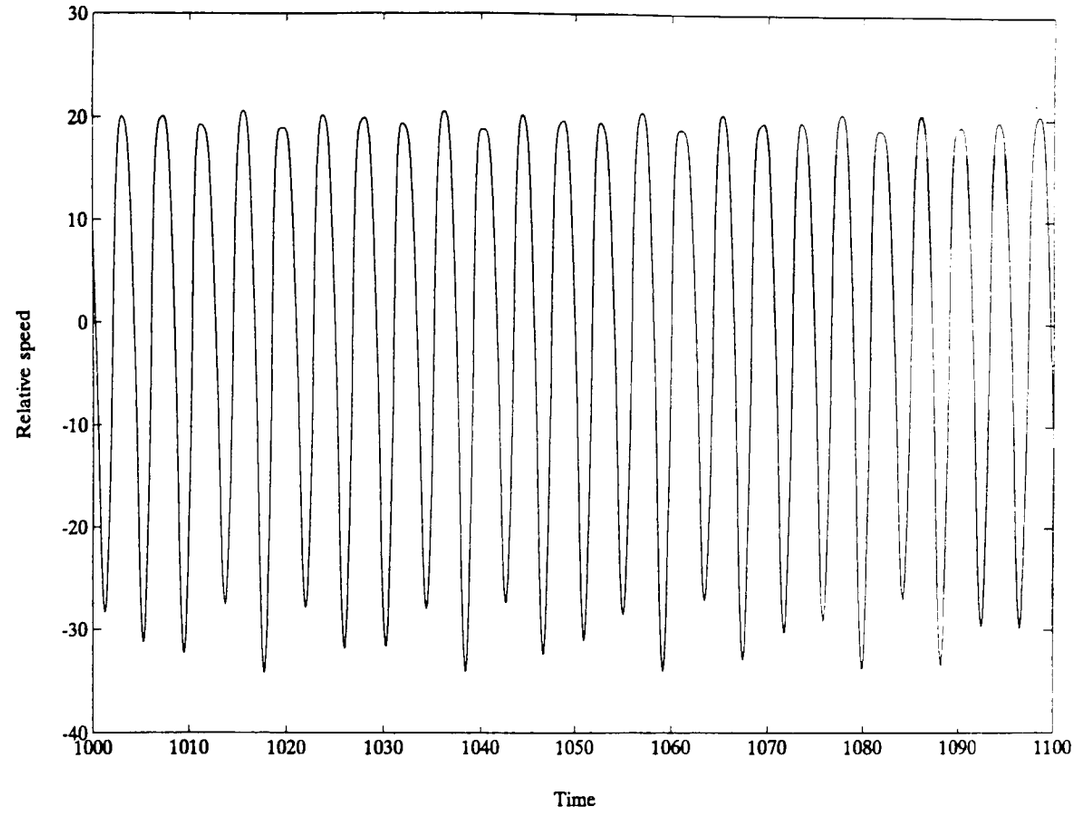


(a)

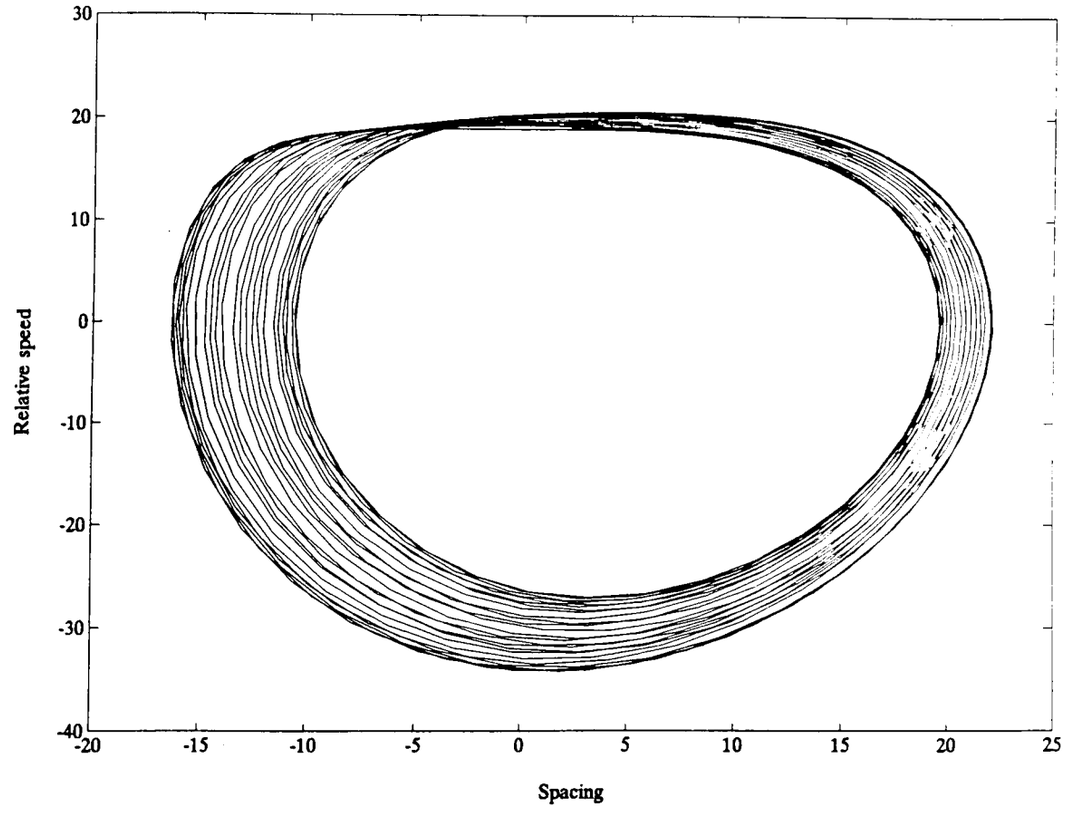


(b)

Figure 4.7 Conveyance of a disturbance along cars in the nonlinear autonomous car-following model, with $N = 10$, $l = 2.3$, $m = 1.1$, $\tau = 1$, $b = 30$, $\bar{u} = 15$. (a) $\alpha = 60$; (b) $\alpha = 100$.

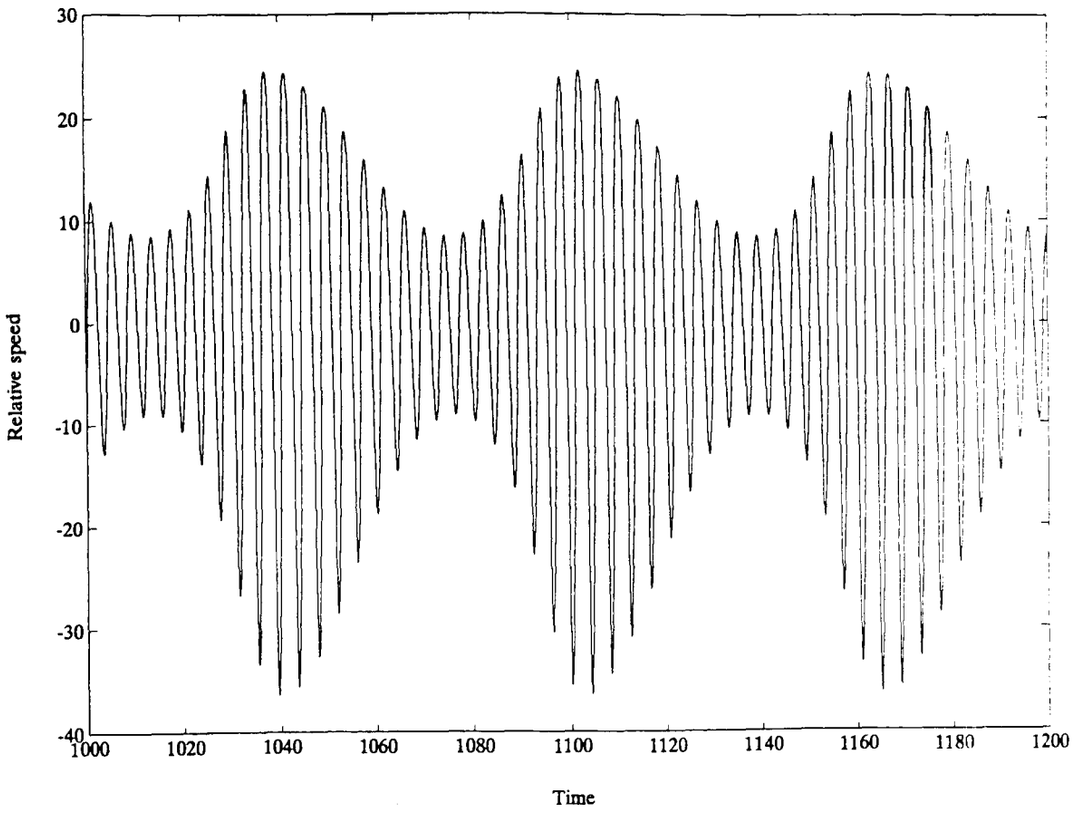


(a)

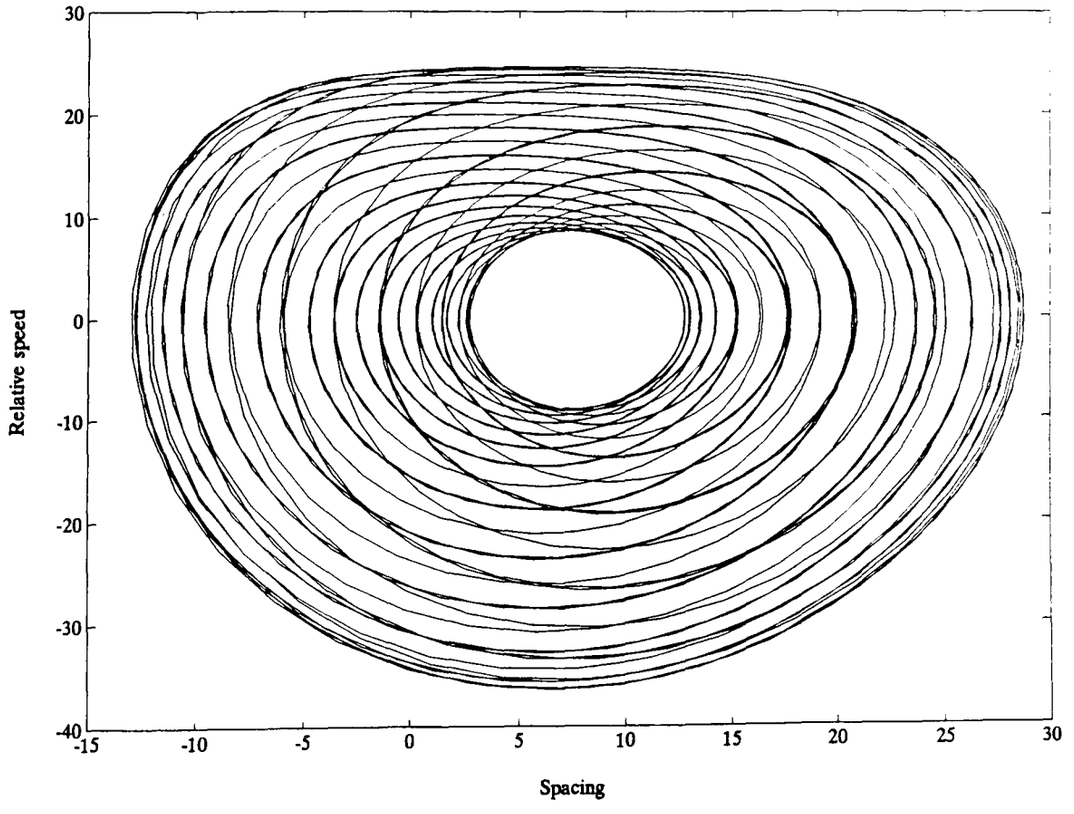


(b)

Figure 4.8 Periodic solution of the nonlinear non-autonomous car-following model, with $N = 2$, $l = 2.0$, $m = 1.0$, $\alpha = 140$, $\tau = 1.0$, $b = 30$, $F = 1.0$, $\rho = 0.9$, $\bar{u} = 20$. (a) Relative speed; (b) relative speed against spacing.

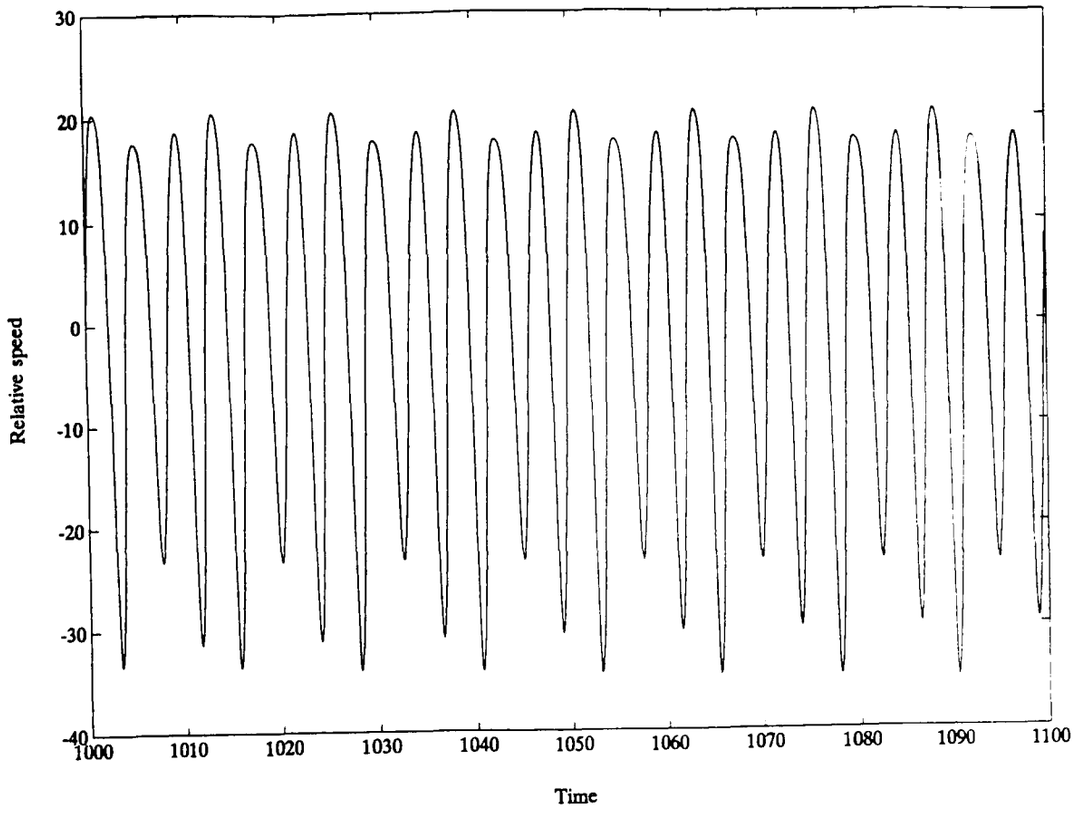


(a)

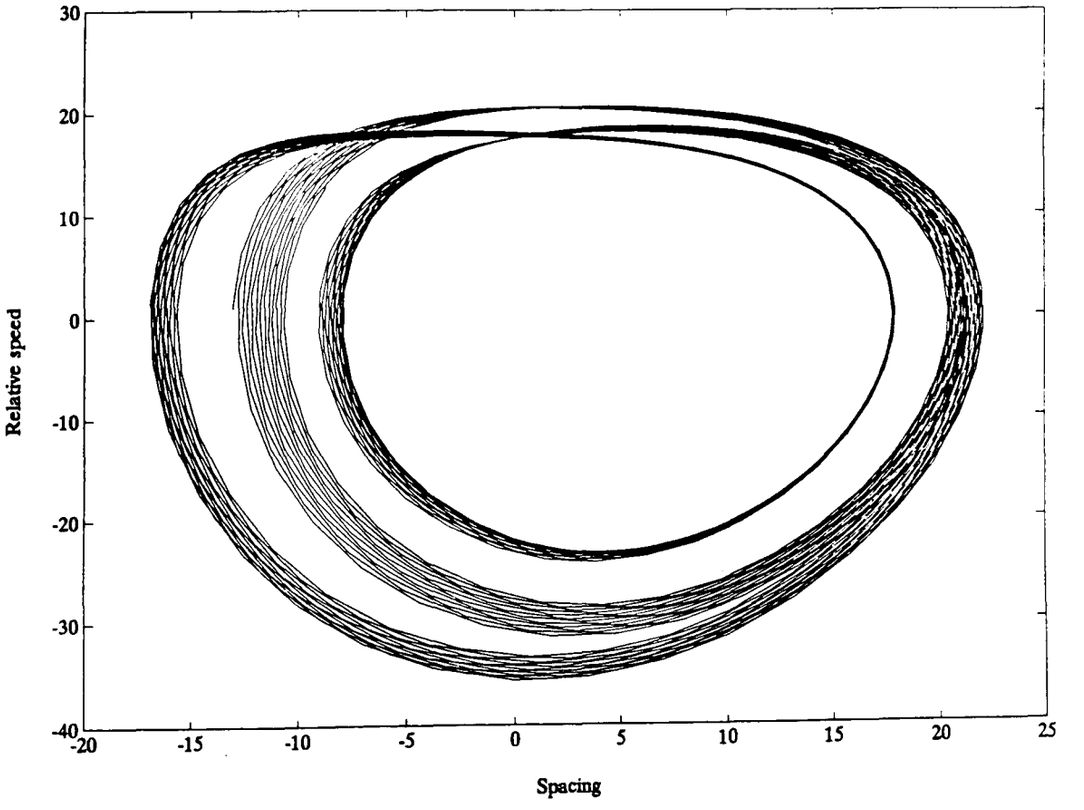


(b)

Figure 4.9 Periodic solution of the nonlinear non-autonomous car-following model, with $N = 2$, $l = 2.0$, $m = 1.0$, $\alpha = 140$, $\tau = 1.0$, $b = 30$, $F = 0.5$, $\rho = 0.1$, $\bar{u} = 20$. (a) Relative speed; (b) relative speed against spacing.

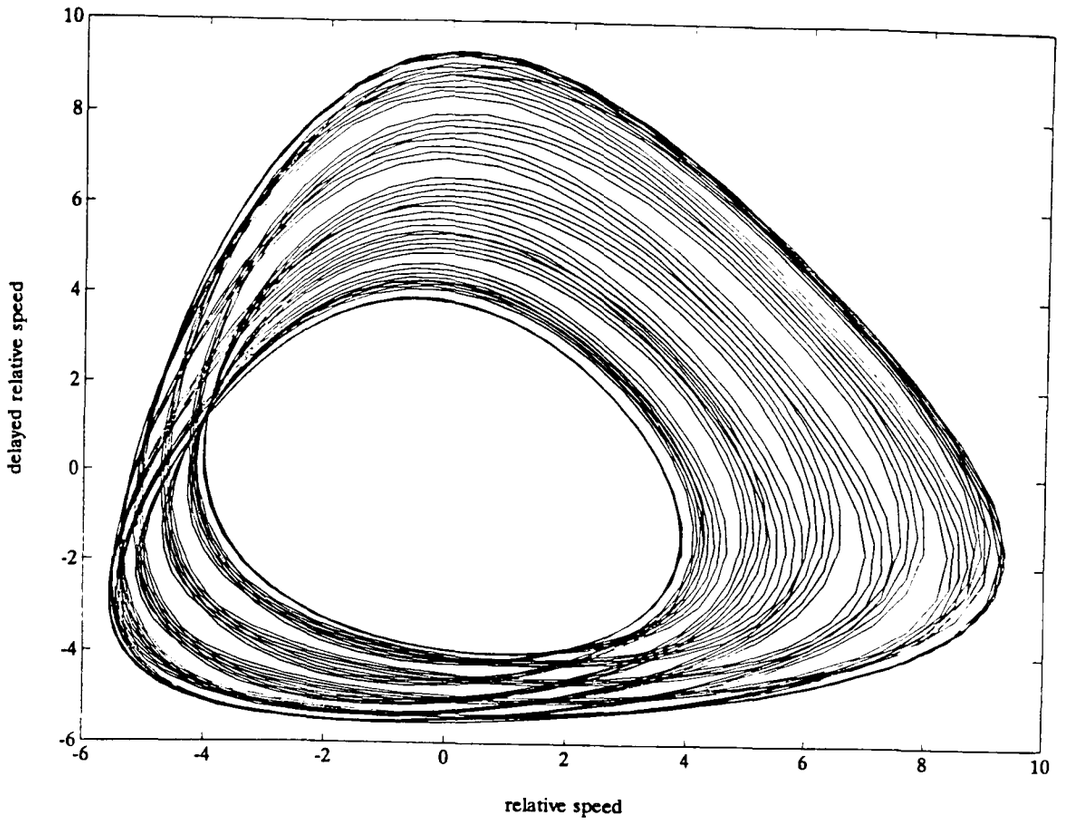


(a)

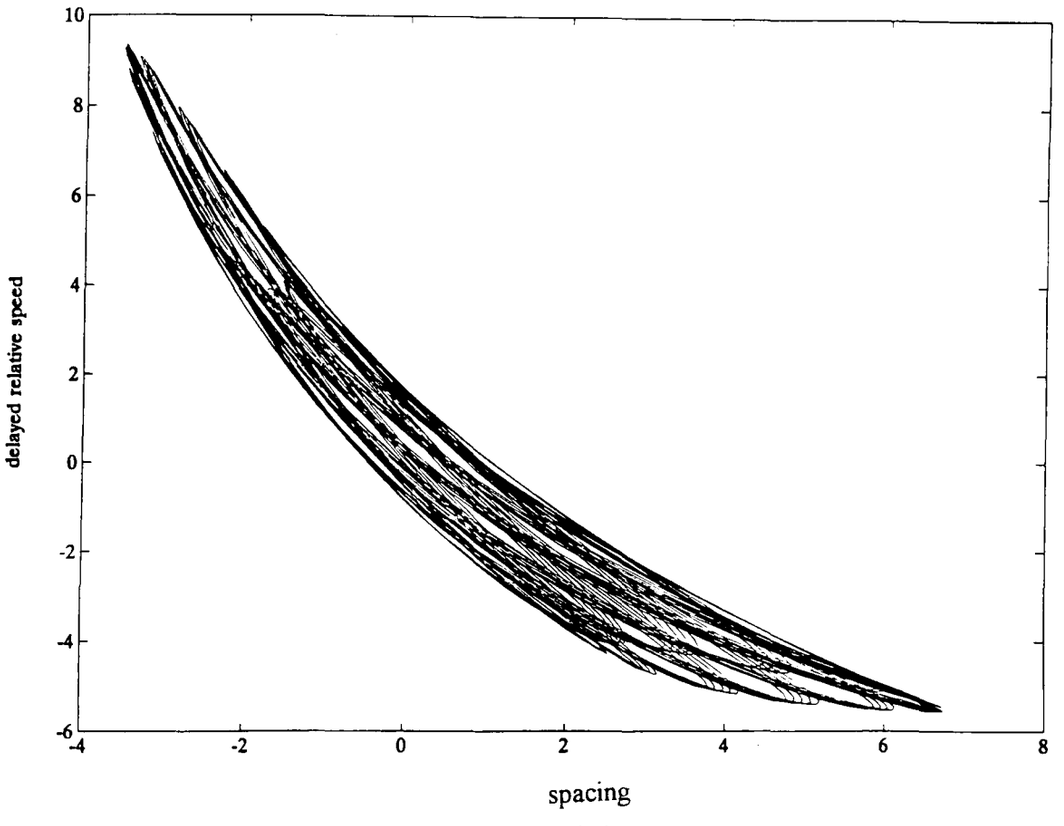


(b)

Figure 4.10 Periodic solution of the nonlinear non-autonomous car-following model, with $N = 4$, $l = 2.5$, $m = 1.3$, $\alpha = 350$, $\tau = 1$, $b = 30$, $F = 1.0$, $\rho = 0.5$, $\bar{u} = 20$. (a) Relative speed of the third car to the fourth car; (b) relative speed against spacing of the third car to the fourth car.

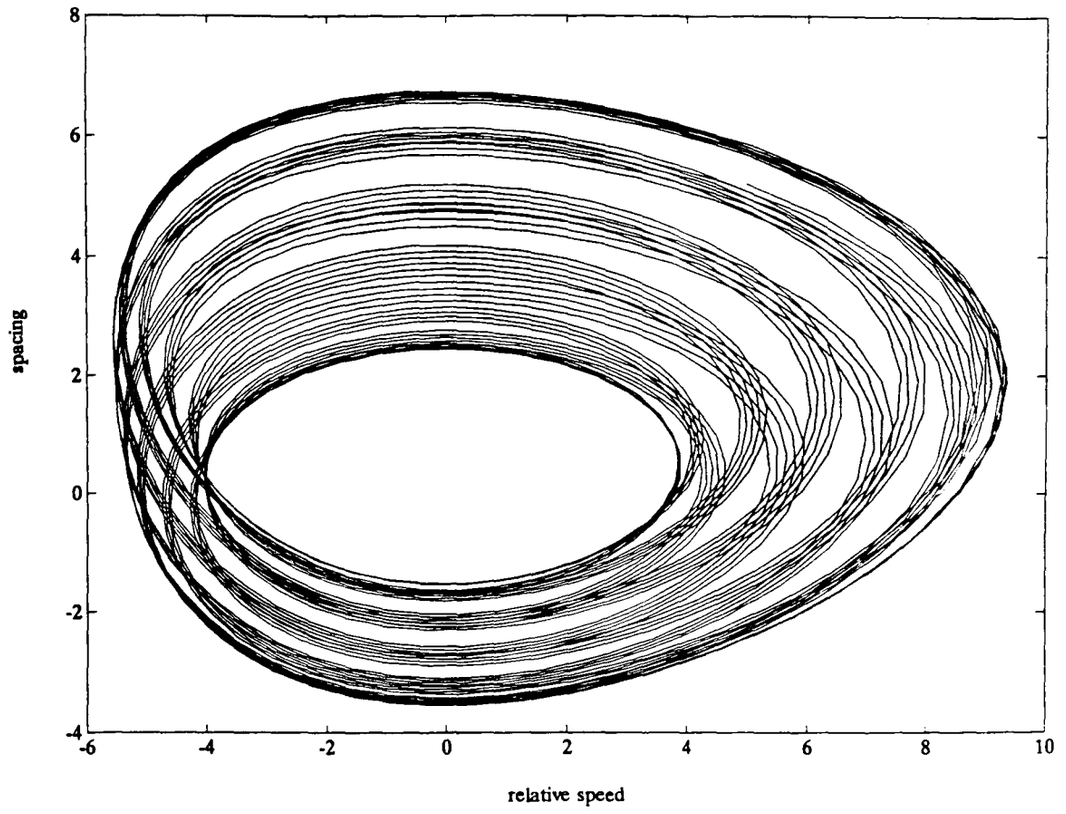


(a)

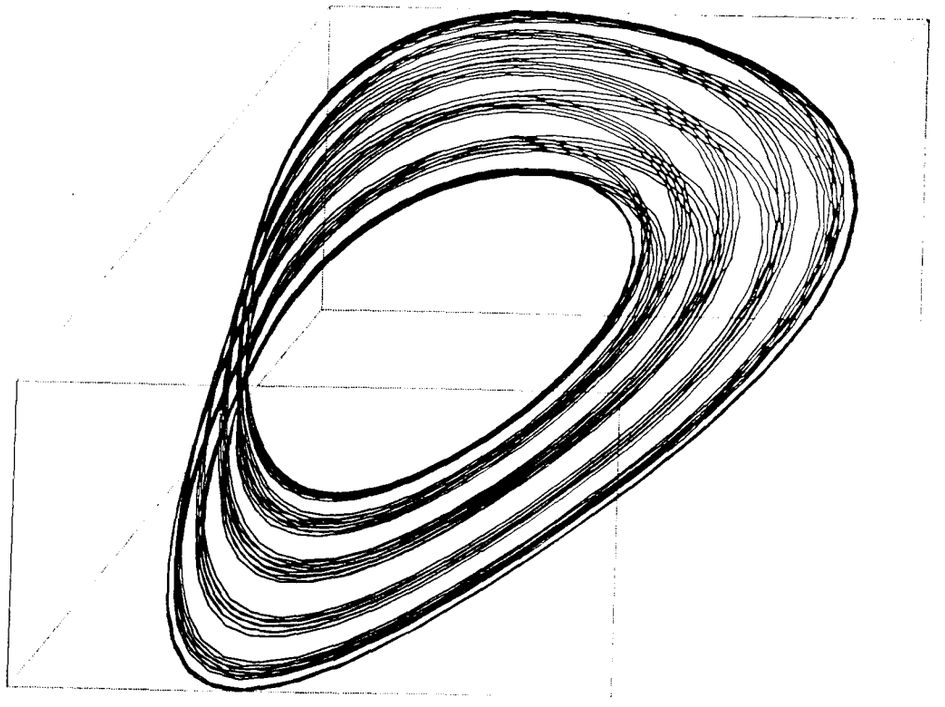


(b)

Figure 4.11 "Quasi-periodic solution" of the nonlinear non-autonomous car-following model, with $N = 4$, $l = 2$, $m = 0$, $\alpha = 165.3$, $\tau = 1$, $b = 10$, $F = 0.5$, $\rho = 1$, $\bar{u} = 15$. (a) Delayed relative speed against relative speed; (b) delayed relative speed against spacing.

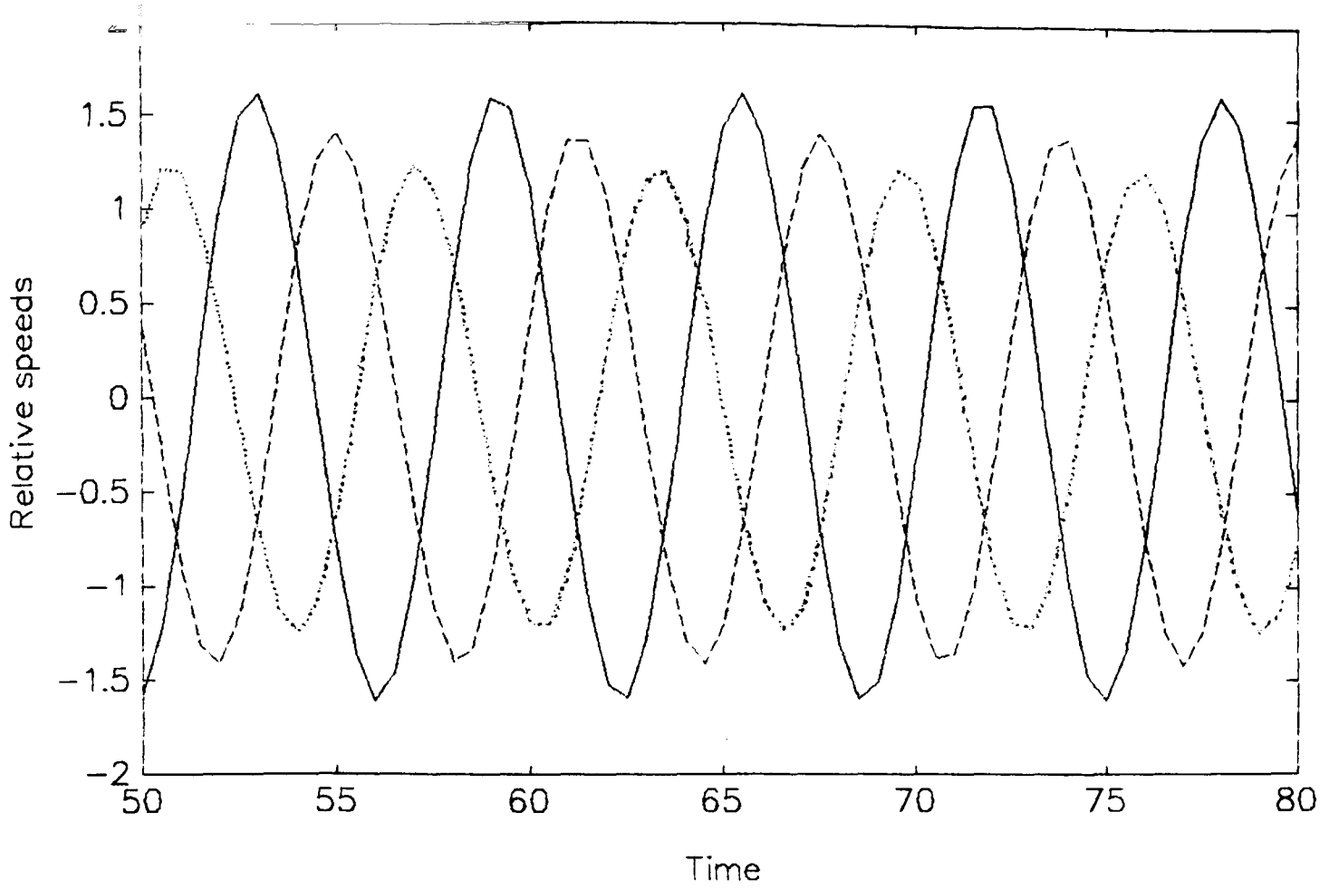


(c)

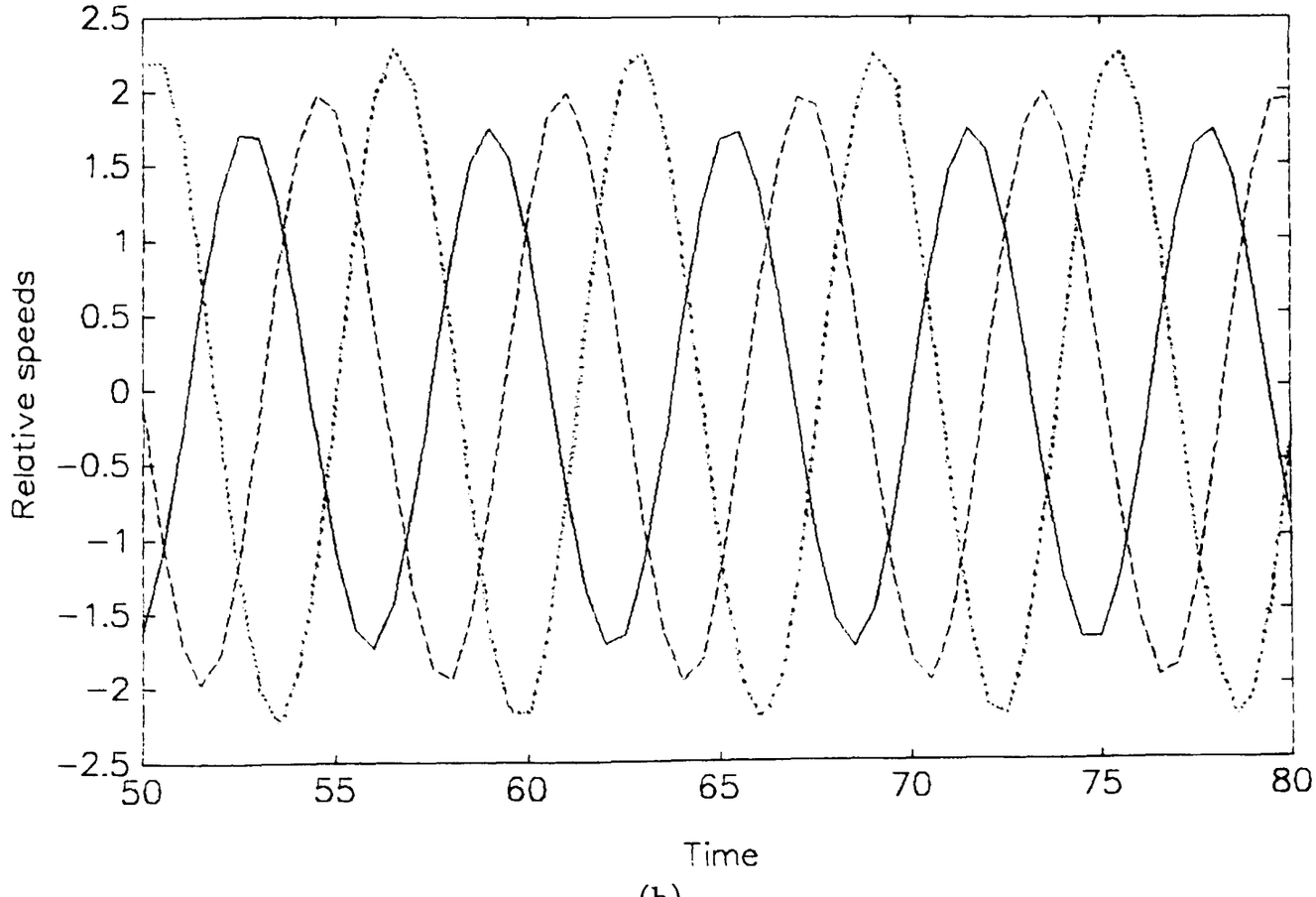


(d)

Figure 4.11 continued. (c) Spacing against relative speed; (d) phase portrait in the 3 dimensional phase space of delayed relative speed, relative speed, and spacing.



(a)



(b)

Figure 4.12 Conveyance of a disturbance along cars in the nonlinear non-autonomous car-following model, with $N = 4$, $l = 2.3$, $m = 1.1$, $b = 30$, $\bar{u} = 15$, $F = 1$, $\rho = 1$, $\tau = 1$. (a) $\alpha = 110$; (b) $\alpha = 130$.

CHAPTER 5. THE DYNAMIC BEHAVIOUR OF THE GRAVITY MODEL

In this chapter, the variations of the flow pattern in an Origin-Destination network (O–D network) are investigated based on the dynamic gravity model for trip distribution. The dynamic gravity model is suggested by Dendrinos and Sonis (1990). To the knowledge of the author of the thesis, there is not any previous study of this model. Here, the model is examined both theoretically and numerically to identify all possible dominant dynamic behaviour in the model.

5.1. INTRODUCTION

The gravity model has been developed for trip distribution, which deals with the problem of determining a non-negative trip matrix $\mathbf{t} = [t_{ij}]$, $i = 1, 2, \ldots, I$, $j = 1, 2, \ldots, J$ satisfying certain marginal constraints. In the dynamic gravity model, it is assumed that the number of trips at each stage, such as each day, each week, and so on, depends on the travel cost at the previous stage, and that the travel cost is a function of the number of trips. Let S be the set of all possible values of $\mathbf{t}$. This will be defined by non-negativity constraints on the t_{ij} and appropriate marginal constraints . Then the dynamic gravity model is defined by a mapping $\mathbf{F}$: $S \to S$,

$$F_{ij}(\mathbf{t}) = \psi_{ij}(\mathbf{t})\, f(c_{ij}(t_{ij})), \quad i = 1, 2, \ldots, I, \ j = 1, 2, \ldots, J. \tag{5.1}$$

Here, $\psi_{ij}(\mathbf{t})$ is an appropriate normalizing factor determined from the marginal constraints, c_{ij} is the travel cost which is normally assumed to be an increasing function of t_{ij}, and $f(\cdot)$ is called the *deterrence function* which relates the number of trips to the travel costs. The map (5.1) defines a discrete-time dynamical system. If n is the discrete time, and $\mathbf{t}(n)$ the O–D flow pattern at time n, then $\mathbf{t}(n+1) = \mathbf{F}(\mathbf{t}(n))$ is the O–D flow pattern at time $n+1$. An equilibrium $\mathbf{t}^e$ (if it exists) in this system is given by $\mathbf{t}^e = \mathbf{F}(\mathbf{t}^e)$. Similarly, a

period-two orbit (if there is one) is defined by two points, say $\mathbf{t}^1$ and $\mathbf{t}^2$, in the phase space such that

$$\mathbf{t}^2 = \mathbf{F}(\mathbf{t}^1), \qquad \mathbf{t}^1 = \mathbf{F}(\mathbf{t}^2).$$

Three types of deterrence function are usually considered (Ortúzar and Willumsen, 1990): (a) exponential function, (b) power function, and (c) combined function. They can be written as:

$$f(c_{ij}) = c_{ij}^{\mu} \exp(-\beta\, c_{ij}),$$

where μ and β are constants. When $\mu = 0$ and $\beta > 0$, f is an exponential deterrence function; when $\mu < 0$ and $\beta = 0$ it is a power deterrence function; and when $\mu > 0$ and $\beta > 0$ it is a combined deterrence function. The exponential and the power deterrence function are both decreasing functions of costs. The combined function is not a monotonic function of the cost; the number of trips increases at first and then decreases with the cost. The analysis of this chapter will show that this difference between the deterrence functions leads to very different kinds of model behaviour.

The normalizing factor $\psi_{ij}(\mathbf{t})$ in (5.1) is chosen so that one or more of the following marginal constraints of an O–D matrix are satisfied

$$\text{(a)} \quad \Sigma_{ij} t_{ij} = 1, \tag{5.2a}$$

$$\text{(b)} \quad \Sigma_j t_{ij} = o_i, \qquad i = 1, 2, \ldots, I, \tag{5.2b}$$

$$\text{(c)} \quad \Sigma_i t_{ij} = d_j, \qquad j = 1, 2, \ldots, J. \tag{5.2c}$$

Replacing $\psi_{ij}(\mathbf{t})$ by appropriate factors we then have three types of models with different constraints, as follows.

(1) Unconstrained model. In this model, only (5.2a) is satisfied and

$$\psi_{ij}(\mathbf{t}) \equiv \psi(\mathbf{t}) = \frac{1}{\Sigma_{kl} f(c_{kl}(t_{kl}))}, \quad i = 1, 2, \ldots I, \; j = 1, 2, \ldots J$$

so that

$$F_{ij}(\mathbf{t}) = \frac{f(c_{ij}(t_{ij}))}{\Sigma_{kl} f(c_{kl}(t_{kl}))}, \quad t_{ij} \geq 0, \ \Sigma_{ij} t_{ij} = 1. \tag{5.3a}$$

In this case, $\psi_{ij}(\mathbf{t})$ does not depend on i or j.

(2) Singly constrained model. There are two kinds of constraints, origin constraint and destination constraint. For an origin-constrained model (5.2b) is met and $\psi_{ij}(\mathbf{t})$ depends on i only. Denote the factor by $a_i(\mathbf{t})$. Then

$$a_i(\mathbf{t}) = o_i \frac{1}{\Sigma_j f(c_{ij}(t_{ij}))}, \quad i = 1, 2, \ldots I,$$

so that

$$F_{ij}(\mathbf{t}) = o_i \frac{f(c_{ij}(t_{ij}))}{\Sigma_l f(c_{il}(t_{il}))}, \quad t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i. \tag{5.3b}$$

While for a destination-constrained model (5.2c) is met and $\psi_{ij}(\mathbf{t})$ depends on j only. Denote the factor by $b_j(\mathbf{t})$. Then

$$b_j(\mathbf{t}) = d_j \frac{1}{\Sigma_i f(c_{ij}(t_{ij}))}, \quad j = 1, 2, \ldots J,$$

so that

$$F_{ij}(\mathbf{t}) = d_j \frac{f(c_{ij}(t_{ij}))}{\Sigma_k f(c_{kj}(t_{kj}))}, \quad t_{ij} \geq 0, \ \Sigma_i t_{ij} = d_j. \tag{5.3c}$$

(3) Doubly constrained model. In this model, both (5.2b) and (5.2c) are satisfied. The normalizing factor is replaced by two sets of constants, $a_i(\mathbf{t})$ and $b_j(\mathbf{t})$, satisfying the equations

$$a_i(\mathbf{t}) = o_i \frac{1}{\Sigma_j b_j(\mathbf{t}) f(c_{ij}(t_{ij}))}, \quad i = 1, 2, \ldots I,$$

and

$$b_j(\mathbf{t}) = d_j \frac{1}{\Sigma_i a_i(\mathbf{t}) f(c_{ij}(t_{ij}))}, \quad j = 1, 2, \ldots J.$$

The model is

$$F_{ij}(\mathbf{t}) = a_i(\mathbf{t})b_j(\mathbf{t})f(c_{ij}(t_{ij})), \quad t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i, \ \Sigma_i t_{ij} = d_j. \tag{5.3d}$$

The factors $a_i(\mathbf{t})$ and $b_j(\mathbf{t})$ are called *balancing factors.* These two sets of factors are inter-dependent: calculation of one set of factors require the values of the other set of factors. A special algorithm is needed in numerical calculations.

The domain or the phase space S is different for the three types of model. The unconstrained gravity model is a map of the set

$$S = \{[t_{ij}]: \ t_{ij} \geq 0, \ \Sigma_{i=1}^{I} \Sigma_{j=1}^{J} t_{ij} = 1 \}$$

into itself. This is the *standard simplex* in IJ dimensional space, and has dimension IJ–1. The origin-constrained model, on the other hand, is a map on a product of sets

$$S = S_1 \times S_2 \times \ldots \times S_I$$

where

$$S_i = \{(t_{i1}, t_{i2}, \ldots, t_{iJ}): \ t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i \}, \quad i=1, 2, \ldots, I,$$

so that

$$S = \{[t_{ij}]: \ t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i, \ i=1, 2, \ldots, I \}.$$

It can be seen that in the origin-constrained model $t_{ij}(n+1)$ depends only on the elements of the ith row of a trip matrix. Therefore the model consists of I independent equations, each on S_i and with a dimension J–1. The set S_i is a simplex in J dimensional space, and can be converted to a standard simplex by dividing each t_{ij} by o_i. Thus each component of the origin-constrained model, after further normalization, is equivalent to the unconstrained model. Similarly, the destination-constrained model consists of J independent equations on

$$S_j = \{(t_{1j}, t_{2j}, \ldots, t_{Ij}): \ t_{ij} \geq 0, \ \Sigma_i t_{ij} = d_j \}$$

and with a dimension of I–1. Again each component of the model can be normalized to get a model equivalent to the unconstrained model.

The doubly constrained model, however, is different from the unconstrained or the singly constrained model. The phase space for the doubly constrained model is

$$S = \{[t_{ij}]: \ t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i, \ i=1, 2, \ldots, I, \ \Sigma_i t_{ij} = d_j, \ j=1, 2, \ldots, J\}$$

This phase space has a dimension of $(I-1)(J-1)$ and is more complicated than that for the unconstrained model. Another complication of the doubly constrained model is, as mentioned earlier, that the two sets of normalizing factors are inter-dependent. This feature can pose difficulties in both theoretical analysis and numerical analysis of the doubly constrained model. Therefore, this model often has to be treated separately from the unconstrained and singly constrained models.

The gravity model (5.1) is examined theoretically in the next section, where the existence, the uniqueness, and stability of an equilibrium in the model are analyzed. Also discussed in the next section is the existence of the user equilibrium and the stochastic user equilibrium in the model. These two equilibria are special flow patterns in a road network used in trip assignment studies (See Chapter 2). In this chapter, the concepts are extended to O–D flow patterns. Numerical analysis is made to find other possible attractors in the model in section 5.3. Liapunov exponents and fractal dimensions are calculated to characterize chaotic attractors found in the model; these are described in sections 5.4 and 5.5 respectively. The chapter is summarized in the last section.

5.2. THEORETICAL ANALYSIS

5.2.1. The existence of an equilibrium

The existence of an equilibrium or a fixed point in (5.1) can be assured by Brouwer's fixed point theorem (Griffel, 1981) which, for the purposes of this thesis, can be stated as follows.

Brouwer's fixed point theorem If Ω is a closed convex set, any continuous map $\mathbf{F}: \Omega \rightarrow \Omega$ has at least one fixed point.

The map defined by (5.1) is continuous and the phase spaces for all three types of

model are closed convex sets. Therefore, in all three cases the map has at least one fixed point in the phase space.

As mentioned in the last section, an equilibrium $\mathbf{t}^e$ in the gravity model (5.1) is given by

$$t_{ij}^e = F_{ij}(t_{ij}^e) = \psi_{ij}(\mathbf{t}^e)\, f(c_{ij}(t_{ij}^e)), \quad i = 1, 2, \ldots, I, \quad j=1, 2, \ldots, J,$$

or

$$\mathbf{t}^e = \mathbf{F}(\mathbf{t}^e).$$

The equilibrium cannot be found analytically unless $\mathbf{F}$ is one dimensional and is such that the above equation is linear or quadratic. However, it can be obtained by numerical calculations.

5.2.2. The uniqueness of the equilibrium

The uniqueness of the equilibrium in (5.1) is examined through a related mathematical programming problem, although other methods of proof may be possible. For the exponential and power deterrence function, the condition for uniqueness is established by showing that the equilibrium coincides with a stationary point which is the unique minimum of a convex objective function. By the same method, a sufficient condition for the uniqueness of the equilibrium is obtained for the model with the combined deterrence function.

The problem to be considered is

$$\text{Minimize } z(\mathbf{t}) = \Sigma_{ij} t_{ij} \ln t_{ij} + \Sigma_{ij} \int_0^{t_{ij}} \left[\beta\, c_{ij}(s) - \mu \ln c_{ij}(s)\right] ds \tag{5.4a}$$

subject to

$$\Sigma_j t_{ij} = o_i \qquad \forall i \tag{5.4b}$$

$$\Sigma_i t_{ij} = d_j \qquad \forall j \tag{5.4c}$$

$$t_{ij} \geq 0 \qquad \forall i,j \tag{5.4d}$$

We have considered the doubly constrained model here. The unconstrained and the singly constrained models can be considered by simply replacing the

constraints (5.4b) and (5.4c) by the appropriate one in (5.2a)–(5.2c). When $t_{ij} = 0$, the expression $t_{ij} \ln t_{ij}$ in the objective function is defined to be zero.

The Lagrangian for the above problem is

$$L(\mathbf{t}, \mathbf{u}, \mathbf{v}) = z(\mathbf{t}) + \Sigma_i u_i (o_i - \Sigma_j t_{ij}) + \Sigma_j v_j (d_j - \Sigma_i t_{ij}),$$

where $\mathbf{u}$ and $\mathbf{v}$ are vectors of Lagrange multipliers. Let $\mathbf{t}^*$ be a stationary point. Then we must have

$$\frac{\partial}{\partial t_{ij}} L(\mathbf{t}^*, \mathbf{u}^*, \mathbf{v}^*) = 0, \quad \forall i,j,$$

or

$$\ln t_{ij}^* + 1 + \beta\, c_{ij}(t_{ij}^*) - \mu \ln c_{ij}(t_{ij}^*) - u_i - v_j = 0, \quad \forall i,j.$$

From this we obtain the stationary point

$$t_{ij}^* = c_{ij}(t_{ij}^*)^{\mu} \exp(-\beta\, c_{ij}(t_{ij}^*)) \exp(-1+u_i) \exp(v_j), \quad \forall i,j.$$

The factors $\exp(-1+u_i) \exp(v_j)$ determined from the constraints (5.4b) and (5.4c) are exactly the same as the normalizing factor in the doubly constrained model. So the stationary point in the programming problem (5.4) is identical to the fixed point in the doubly constrained model (5.3d). If the objective function is strictly convex at $\mathbf{t}^*$, then $\mathbf{t}^*$ is a local minimum. If the objective function is convex elsewhere in the feasible region defined by the constraints (5.4b)–(5.4d) and if the region is convex as well, then $\mathbf{t}^*$ is a global, or unique, minimum in the region, and so is the equilibrium in the phase space of dynamic gravity model. Now the feasible region is convex since it is defined by linear equalities and inequalities. Also it follows from Theorem 1 in Evans (1973) that the first term of the objective function is strictly convex in the feasible region. It remains to be proved that the second term of the function is strictly convex at $\mathbf{t}^*$ and is convex elsewhere. Denote this term by z_2. Then

$$z_2(\mathbf{t}) = \Sigma_{ij} \int_0^{t_{ij}} \left[\beta\, c_{ij}(s) - \mu \ln c_{ij}(s)\right] \mathrm{d}s.$$

The elements of the Hessian matrix for this function are given by

$$\frac{\partial}{\partial t_{kl}}\left(\frac{\partial z_2(\mathbf{t})}{\partial t_{ij}}\right) = \begin{cases} c_{ij}'(t_{ij})\left[\beta - \dfrac{\mu}{c_{ij}(t_{ij})}\right], & \text{for} \quad kl=ij \\ 0, & \text{otherwise} \end{cases}$$

This means that the Hessian is a diagonal matrix. The condition for the objective function to be (strictly) convex is that the derivatives are (strictly) positive so that the Hessian is positive semi-definite (positive definite), that is

$$c_{ij}'(t_{ij}^*)\left[\beta - \frac{\mu}{c_{ij}(t_{ij}^*)}\right] > 0, \quad \text{and} \quad c_{ij}'(t_{ij})\left[\beta - \frac{\mu}{c_{ij}(t_{ij})}\right] \geq 0, \quad \forall t_{ij}$$

Since c_{ij} is an increasing function of t_{ij}, the above conditions can be reduced to

$$\left[\beta - \frac{\mu}{c_{ij}(t_{ij}^*)}\right] > 0, \quad \text{and} \quad \left[\beta - \frac{\mu}{c_{ij}(t_{ij})}\right] \geq 0, \ \forall t_{ij} \tag{5.5}$$

These are automatically satisfied if the exponential or the power deterrence function is used for in the former $\beta > 0$, $\mu = 0$, and in the latter $\beta = 0$, $\mu < 0$. In both cases, there is equality only when $t_{ij} = 0$, that is, at the boundary of the feasible region. Therefore, when the exponential or the power deterrence function is used, the stationary point is a unique minimum in the feasible region and the equilibrium in the gravity model is unique. With the combined function, if the values of parameters are such that the above conditions are satisfied then the optimum solution and so the equilibrium is unique. Otherwise, the solution and the equilibrium is not necessarily unique. An example in which the equilibrium is not unique in the model with the combined deterrence function will be shown in the numerical analysis.

5.2.3. The stochastic user equilibrium

In this section we will show that the unique equilibrium in the gravity model with the exponential deterrence function is equivalent to the stochastic user equilibrium (SUE), with the user equilibrium (UE) as its special case. The UE and the SUE are two notions of flow patterns in trip assignment; these have been introduced in Chapter 2 and will be described in detail in Chapter 6. Here, the extension of the idea to trip distribution is outlined.

The dynamic process of trip distribution is the same as that of trip assignment in that they both deal with problems of selection among alternatives, although in the former the alternatives are O–D pairs and in the latter they are the routes. Therefore, the notions of the UE and SUE in trip assignment models can be used to describe the equilibria in trip distribution models.

In trip distribution, it is plausible to assume that trip makers choose the O–D pair in an area with the minimum cost. The cost between each O–D pair varies with the flow between the O–D pair; the cheapest O–D pair may become more expensive because of congestions. In addition, the appropriate marginal constraints of trip matrix must be satisfied. Consequently, trips in the area may be spread among more than one O–D pair. A flow pattern may be reached when the cost of all trip makers is the minimum and no trip maker can reduce the cost by changing to another O–D pair. This flow pattern is the *user equilibrium in an O–D network.* In this UE definition, however, it is assumed implicitly that trip makers know the travel cost between every O–D pair and they are identical in their choice behaviour. This presumption can be relaxed by introducing a *perceived travel cost* between O–D pairs, which can reflect the difference or the randomness among trip makers in their choice behaviour. Each trip maker may perceive a different travel cost for the same O–D pair. An equilibrium will be reached when the cost of all trip makers is *thought* to be the minimum and no trip maker can reduce the perceived cost by changing to another O–D pair. This is the *stochastic user equilibrium in an O–D network.* Note that this equilibrium is called stochastic user equilibrium simply because the variations in drivers' perception of travel cost is included. This is where the user equilibrium differs from the stochastic user equilibrium. The flow variables used to define the stochastic user equilibrium, however, is deterministic variables rather than stochastic ones.

In trip assignment studies in the literature, the problem of finding the UE or SUE has been formulated as equivalent mathematical programming problems. For trip distribution, we can have similar results. Consider the program (5.4) with $\mu = 0$

$$\text{Minimize } z_3(\mathbf{t}) = \Sigma_{ij} t_{ij} \ln t_{ij} + \beta \Sigma_{ij} \int_0^{t_{ij}} c_{ij}(s)\, \mathrm{d}s,$$

subject to

$$\begin{aligned} \Sigma_j t_{ij} &= o_i & \forall i \\ \Sigma_i t_{ij} &= d_j & \forall j \\ t_{ij} &\geq 0 & \forall i,j. \end{aligned}$$

Clearly, the solution of the program is the same as the equilibrium of the dynamic gravity model with the exponential deterrence function. The objective function of the program has the same form as that of the mathematical programm formulated by Fisk (1980) to solve the logit-based SUE traffic assignment (See Chapter 6). In that objective function, the variables are the route flows and link flows on a road network, while here they are O–D flows of an O–D network. The unique minimum point of that objective function gives SUE for trip assignment; the unique optimum here is the SUE for an O–D network, which is also the unique equilibrium in the dynamic gravity model with the exponential deterrence function.

The above objective function consists of two terms with a weighting coefficient β. The first term is the entropy of the O–D matrix. If the costs are constants and do not depend on the flow, the entropy of the O–D matrix is maximized with the cost constraint (Ortúzar and Willumsen, 1990), and the solution is the same as that given by the static classical gravity model (the model obtained by omitting the time n in the dynamic gravity model with the exponential deterrence function. See, for example, Ortúzar and Willumsen, 1990). The second term, though does not seem to have a clear physical meaning, is the same as the objective function of the program formulated by Beckmann *et al.* (1956) to find the UE for trip assignment (Again, see Chapter 6). Thus, if β is very large, the above objective function tends to the second term and the solution tends to the UE for an O–D network.

Therefore, the equilibrium in the dynamic gravity model is the stochastic user equilibrium, in which both the flow pattern given by the static gravity model with the exponential deterrence function and the user equilibrium are special cases.

5.2.4. The Stability of the equilibrium

Generally speaking, an equilibrium in a dynamical system is not necessarily stable. The stability normally depends on the values of parameters in the model.

The stability of the gravity model is examined in this section.

A general analysis of the stability and the bifurcation of the gravity model is difficult, especially for the doubly constrained model. Therefore, here, the stability of equilibria is investigated for the unconstrained and singly constrained models only. The doubly constrained model will be considered in the numerical analysis.

The unconstrained and singly constrained models (5.3a) – (5.3c) can be written in the more general form, using single subscripts for simplification,

$$F_i(\mathbf{t}) = \frac{f(c_i(t_i))}{\Sigma_j f(c_j(t_j))}, \quad i = 1, 2, \ldots, K, \tag{5.6}$$

where

$$f(c_i) = c_i^{\mu} \exp(-\beta c_i).$$

When $K = IJ$ the equation represents an unconstrained gravity model; while when K equals I or J the equation represents one component of a singly constrained model with o_i or d_j being set to 1 for normalization.

We will consider first a one dimensional model, and then move onto multi-dimensional models. A one dimensional model can occur where there are one origin and two destinations, or vice versa. The model is

$$F_i(\mathbf{t}) = \frac{f(c_i(t_i))}{f(c_1(t_1)) + f(c_2(t_2))}, \quad i = 1, 2. \tag{5.7}$$

A fixed point is locally asymptotically stable if the derivative at the point is within (–1, 1) (Parker and Chua, 1989). From (5.7), we have

$$\frac{\mathrm{d}}{\mathrm{d}t_1} F_1(\mathbf{t}) = \frac{f(c_1(t_1))\ f(c_2(t_2))}{[f(c_1(t_1)) + f(c_2(t_2))]^2} \left[c_1'(t_1) \left[\frac{\mu}{c_1(t_1)} - \beta \right] \right.$$

$$\left. + c_2'(t_2) \left[\frac{\mu}{c_2(t_2)} - \beta \right] \right], \tag{5.8}$$

where the primes stand for derivatives and t_2=1−t_1. At the fixed point $\mathbf{t}^e$,

$$\left|\frac{\mathrm{d}}{\mathrm{d}t_1}F_1(\mathbf{t}^e)\right| = t_1^e\, t_2^e\left|c_1'(t_1^e)\left[\frac{\mu}{c_1(t_1^e)}-\beta\right]+c_2'(t_2^e)\left[\frac{\mu}{c_2(t_2^e)}-\beta\right]\right|$$

The equilibrium is stable if

$$t_1^e\, t_2^e\left|c_1'(t_1^e)\left[\frac{\mu}{c_1(t_1^e)}-\beta\right]+c_2'(t_2^e)\left[\frac{\mu}{c_2(t_2^e)}-\beta\right]\right| < 1. \tag{5.9}$$

Since $t_1^e + t_2^e = 1$, we have $t_1^e\, t_2^e \leq 0.25$. Therefore,

$$\left|\frac{\mathrm{d}}{\mathrm{d}t_1}F_1(\mathbf{t}^e)\right| \leq 0.25\left|c_1'(t_1^e)\left[\frac{\mu}{c_1(t_1^e)}-\beta\right]+c_2'(t_2^e)\left[\frac{\mu}{c_2(t_2^e)}-\beta\right]\right|$$

So, a sufficient condition for stability of the equilibrium can be given by

$$\left|c_1'(t_1^e)\left[\frac{\mu}{c_1(t_1^e)}-\beta\right]+c_2'(t_2^e)\left[\frac{\mu}{c_2(t_2^e)}-\beta\right]\right| < 4.0. \tag{5.10}$$

For the power deterrence function, $\beta = 0$ and $\mu < 0$, the stability condition reduces to

$$-\mu\left[\frac{c_1'(t_1^e)}{c_1(t_1^e)}+\frac{c_2'(t_2^e)}{c_2(t_2^e)}\right] < 4.0, \tag{5.10a}$$

while for the exponential deterrence function, $\beta > 0$ and $\mu = 0$, the condition reduces to

$$\beta\left[c_1'(t_1^e) + c_2'(t_2^e)\right] < 4.0. \tag{5.10b}$$

In both cases, the equilibrium is unique. It can be seen that when the power or the exponential deterrence function is used, the equilibrium is more likely to be stable if μ or β is smaller, although the stability depends on the cost and their derivatives at the equilibrium as well. With combined deterrence function, however, the stability of the equilibrium depends on the relative magnitude of μ and β, as well as the costs and their derivatives.

It can be observed from (5.8) that if the deterrence function is exponential or power function, the derivative is negative in the whole phase space, which means

that $F_1(t_1, t_2) = F_1(t_1, 1-t_1)$ in (5.7) is a decreasing function of t_1. It can be easily shown (Zhang, 1994) that if a one dimensional map is a decreasing map, there are only two possible stable steady states: fixed points and period–2 orbits. If the condition (5.9) is violated, as a result of changes of values of parameters, the equilibrium will become a period-two orbit. Therefore, in the one-dimensional model with the exponential or the power deterrence function the fixed point and period-two orbits are the only possible steady states; trajectories starting from any initial conditions in the phase space approach one of the steady states.

In a multi-dimensional model the local stability of an equilibrium can be determined by the eigenvalues of the Jacobian matrix at the equilibrium. An equilibrium is locally asymptotically stable if the magnitude of all the eigenvalues is less than 1 (Parker and Chua, 1989). Since it is not possible to evaluate the eigenvalues of the Jacobian matrix of $\mathbf{F}$ analytically, a sufficient condition for the stability of the equilibrium can be given by bounding the eigenvalues of the Jacobian matrix. There are several ways of bounding eigenvalues of a matrix. Here the norm of a matrix is used as a bound, based on the following theorem (Lancaster, 1969, page 201).

Theorem If $\mathbf{A} \in \mathbb{R}_{n\times n}$ and $\lambda_{\mathbf{A}} = \max_i |\lambda_i|$ $(1 \le i \le n)$ then for any matrix norm, $\lambda_{\mathbf{A}} \le \|\mathbf{A}\|$.

It is desired that the type of norm to be used is small in magnitude and has a simple form. Based on these considerations, the p–norm is used with $p = 1$.

The partial derivatives of $\mathbf{F}$ at the equilibrium in (5.6) are:

$$\frac{\partial}{\partial t_j} F_i(\mathbf{t}^e) = t_j^e (\delta_{ij} - t_i^e)\, c_j'(t_j^e) \left[\frac{\mu}{c_j(t_j^e)} - \beta\right],$$

where

$$\delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j \end{cases}.$$

Let $\mathbf{J}$ denote the Jacobian matrix of $\mathbf{F}$ at the equilibrium, with elements $\frac{\partial}{\partial t_j} F_i(\mathbf{t}^e)$. Then the 1–norm of this matrix is

$$\|\mathbf{J}\|_1 = \max_j \Sigma_i \left| \frac{\partial}{\partial t_j} F_i(\mathbf{t}^e) \right|$$

$$= \max_j \left[2\ t_j^e\ (1-t_j^e)\ c_j'(t_j^e) \left| \frac{\mu}{c_j(t_j^e)} - \beta \right| \right]$$

$$= 2\ t_{j^m}^e\ (1-t_{j^m}^e)\ c_{j^m}'(t_{j^m}^e) \left| \frac{\mu}{c_{j^m}(t_{j^m}^e)} - \beta \right|,$$

where j^m is the index at which the sum is maximum. Again, since

$$t_{j^m}^e\ (1-t_{j^m}^e) \leq 0.25,$$

we have

$$\left\| \left[\frac{\partial}{\partial t_j} F_i(\mathbf{t}^e) \right] \right\|_1 \leq 0.5\ c_{j^m}'(t_{j^m}^e) \left| \frac{\mu}{c_{j^m}(t_{j^m}^e)} - \beta \right|.$$

According to the above theorem, no eigenvalue of the Jacobian matrix is bigger than the 1–norm. Therefore,

$$c_{j^m}'(t_{j^m}^e) \left| \frac{\mu}{c_{j^m}(t_{j^m}^e)} - \beta \right| < 1/0.5 = 2 \tag{5.11}$$

can serve as a sufficient condition for the stability of the equilibrium. It can be observed that this condition bears some resemblance to (5.10), the condition for the stability of the equilibrium in the one dimensional model. Again, with the power and the exponential deterrence function, the equilibrium is unique. For the power deterrence function, $\beta = 0$ and $\mu < 0$, (5.11) reduces to

$$-\mu \frac{c_{j^m}'(t_{j^m}^e)}{c_{j^m}(t_{j^m}^e)} < 2, \tag{5.11a}$$

while for the exponential deterrence function, $\beta > 0$ and $\mu = 0$, the condition becomes

$$\beta\ c_{j^m}'(t_{j^m}^e) < 2. \tag{5.11b}$$

The stability of the equilibrium in the gravity models depends on the values of parameters μ, β and the cost at the equilibrium in the similar way to that for the one dimensional model. With the power and the exponential deterrence functions, the smaller μ or β is, the more likely that the equilibrium is stable. With the combined deterrence function, however, the stability depends on both μ and β. These points, together with the effects of parameters in the cost function will be considered in the numerical analysis.

Further analysis of the gravity model cannot be made theoretically; numerical analysis is necessary to identify other possible steady-state behaviour in higher dimensional models and in the doubly constrained model.

5.3. NUMERICAL ANALYSIS

In this section, the gravity model (5.1) is investigated numerically to identify other possible attractors, besides the point attractor mentioned in the theoretical analysis. Numerical calculations involve iterations of the equation $\mathbf{t}(n+1) = \mathbf{F}(\mathbf{t}(n))$ defined by (5.1) for given initial conditions and values of parameters. The iteration is made until a steady state, or an attractor, is reached. If an attractor is more complicated than a fixed point or a period-two orbit, it can be examined by, for example, time series and phase portrait plotting. It can also be examined by spectral analysis when it is necessary to check the periodicity of the attractor. The behaviour of a dynamical system can be best shown by bifurcation diagrams. These are diagrams of steady states of a system against the value of a parameter, or the bifurcation parameter. The bifurcation diagrams are produced by increasing the bifurcation parameter step by step and obtaining the attractors at each step by iterating the model. Only steady states are plotted; transients are removed.

The cost function $c_{ij}(t_{ij})$ used for the numerical analysis is (2.10b) (Dendrinos and Sonis, 1990), which, by the notations here, can be written as

$$c_{ij}(t_{ij}) = c^0_{ij}\left[1 + \alpha\left[\frac{t_{ij}}{q_{ij}}\right]^{\gamma}\right]. \tag{5.12}$$

The model (5.1) is investigated with different numbers of origins and

destinations, using each of the three forms of the deterrence function. Different initial conditions and different values of parameters are tried. The unconstrained and singly constrained models are described in the next subsection, which is followed by the descriptions of the doubly constrained model.

5.3.1. The Unconstrained or singly constrained model

The analysis of the unconstrained and the singly constrained models is based on the generalized form (5.6) of the two models:

$$F_i(\mathbf{t}) = \frac{f(c_i(t_i))}{\Sigma_j f(c_j(t_j))}, \quad i = 1, 2, \ldots, K,$$

where

$$f(c_i) = c_i^{\mu} \exp(-\beta c_i).$$

Using single subscripts, the cost function (5.12) can be written as

$$c_i(t_i) = c_i^0 \left[1 + \alpha \left[\frac{t_i}{q_i}\right]^{\gamma}\right].$$

Since it has been found that models with the power or the exponential deterrence function have very different behaviour from those with the combined deterrence function, they will be described separately.

Models with the power or the exponential deterrence function

In section 5.2.4. it has been shown that a one dimensional model with the exponential or the power deterrence function could only have point attractors and period-two attractors. Numerical calculations seem to show that this is true for higher dimensional models as well. When the values of parameters are small, a stable equilibrium is always approached. When they become larger, the solutions become a stable period-two orbit. Figure 5.1 is a bifurcation diagram for α in a model of two origins and two destinations with $\mathbf{c}^0 = (c_1^0\ c_2^0\ c_3^0\ c_4^0) = (1.4\ 1.2\ 1.8\ 1.6)$ and $\mathbf{q} = (q_1\ q_2\ q_3\ q_4) = (0.17\ 0.15\ 0.25\ 0.23)$. For each value of parameter, the state of a single point means a point attractor and the state of

two points a period-two attractor. Bifurcation diagrams for all other parameters (γ, β, and μ) are also produced and they all look very much the same as Figure 5.1.

The above behaviour in the gravity model is not unexpected. The trip rates at each iteration depend on the costs at the previous iteration by the deterrence function (with the parameter β or μ), while the cost depends on the trip rate by the cost function (with the parameters α and γ). The former is a decreasing function and the latter an increasing one. Therefore, a reduction in trips between an O–D pair at one iteration will cause the cost to decrease. This O–D pair will attract more trips at the next iteration. When the values of parameters are small enough, certain changes of flow in one iteration will cause smaller changes of costs and so flows at the next iteration and the process may converge to an equilibrium. When the values of parameters are large, however, changes of flows may become larger and larger at each iteration and may cause oscillations.

Models with the combined deterrence function

When the combined deterrence function is used and when the number of dimensions is 1 or 2, again, only fixed points and period-two orbits are found. When the dimension is higher (3 or more), however, more complicated behavior occurs in the model. Period doubling and apparently irregular behavior or chaos are found to be quite typical. Results of calculations of a model with two origins and two destinations will be shown first. The same values of $\mathbf{c}^0$ and $\mathbf{q}$ as those in Figure 5.1 are used. Figure 5.2 shows one of the chaotic attractors found in the model. Three projections of the phase portrait (Figures 5.2a–5.2c) are plotted here, which show that the attractor is geometrically a very complicated object. The power spectrum (Figure 5.2d) is continuous, indicating stochastic behavior. When the initial condition is changed slightly, the orbit will soon diverge. The *sensitive dependence on the initial conditions* is shown in Figure 5.2e, where the solid line is the series starting at $\mathbf{t} = [0.0300 \;\; 0.3521 \;\; 0.5313 \;\; 0.0866]$, while the dashed line is the series starting at $\mathbf{t} = [0.0301 \;\; 0.3520 \;\; 0.5313 \;\; 0.0866]$. The starting time is n=1000. It can be seen that the orbits distinguish themselves after less than 50 iterations.

To show more features of the dynamic behaviour in the model, bifurcation diagrams were produced, using all parameters as bifurcation parameters. When

producing these diagrams, the initial conditions were taken in two ways. One way is simply to use the same initial values for all steps of the parameter. There may be a long transient to remove in this way. The other way is to use the final states of the previous step of the parameter. Transients can thus be shortened. By starting from different points, different attractors may be detected if there is more than one attractor for the same values of the parameters.

Figures 5.3–5.4 are two sets of bifurcation diagrams for β but with different values of μ. Plotted on the diagrams are the number of trips from origin 1 to destination 1. In both cases, different starting points lead to different bifurcation sequences and different sets of attractors, implying that there is more than one attractor for the same value of β. Some of these diagrams are a little coarse. It needs to be pointed out that these diagrams use large amounts of computer time to produce. Figures 5.5–5.6 are local enlargements of Figures 5.3–5.4, respectively, showing the bifurcation sequences in more detail.

Several features can be seen in these diagrams. First, as values of parameters vary, the steady states often change from fixed points to period-two orbits, and then period 4, and then period 8, etc. This is a *period doubling* sequence which is a typical route to chaos (Thompson and Stewart, 1988). However, not all periodic doubling leads to chaos; there may be only several doublings, followed for example by period undoubling or by other bifurcations. What is more, chaotic behaviour is often followed by periodic undoubling sequences as well. For example, in Figure 5.5a, where β is the bifurcation parameter, the period doubling leads to chaos, followed by periodic undoubling. On the other hand, in Figure 5.5b, there is the period doubling up to 16 and then undoubling to period 2 without going through chaos. Secondly, there are some *periodic windows*, or stable periodic orbits in the chaotic regimes. See, for example, Figure 5.6a. Thirdly, there are many discontinuous points in these bifurcation diagrams. This may because more than one attractor coexists in the phase space for the same value of parameter. This is confirmed by iterations starting from different initial conditions with the same values of parameters. Different attractors are approached from different initial conditions, producing different bifurcation diagrams.

In section 5.2.2, we concluded that with the combined deterrence function, the equilibrium in the model is not necessarily unique. Here we have an example for this conclusion. In Figures 5.3a and 5.3b when β is between 1 and 1.5, we

have different equilibria from different initial conditions.

Similar calculations are made with other parameters, μ, α, and γ, and to models with more origins and destinations. Similar behavior to those described above are found. It can be seen that the unconstrained or singly constrained models with the combined deterrence function exhibit a very rich dynamic behaviour. The chaotic behaviour will be examined further in section 5.4.

5.3.2. The doubly constrained model

In this subsection, the doubly constrained model (5.3d) with the cost function (5.12) is examined numerically. The model is

$$F_{ij}(\mathbf{t}) = a_i(\mathbf{t})b_j(\mathbf{t})f(c_{ij}(t_{ij})), \quad t_{ij}\geq 0, \ \Sigma_j t_{ij} = o_i, \ \Sigma_i t_{ij} = d_j \tag{5.13a}$$

where the $a_i(\mathbf{t})$ and $b_j(\mathbf{t})$ satisfy the equations

$$a_i(\mathbf{t}) = o_i \frac{1}{\Sigma_j b_j(\mathbf{t}) f(c_{ij}(t_{ij}))}, \quad i = 1, 2, \ldots I, \tag{5.13b}$$

and

$$b_j(\mathbf{t}) = d_j \frac{1}{\Sigma_i a_i(\mathbf{t}) f(c_{ij}(t_{ij}))}, \quad j = 1, 2, \ldots J. \tag{5.13c}$$

This model cannot be iterated directly like the unconstrained or singly constrained model, because it contains two sets of parameters $a_i(\mathbf{t})$ and $b_j(\mathbf{t})$ which are interdependent. The calculation of one set needs the values of the other set. This suggests an iteration process. The method from Ortúzar and Willumsen (1990) will be used here. Given the values of deterrence functions for each O–D pair, $f(c_{ij})$, the algorithm in outline is as follows:

(1) Set all $b_j(\mathbf{t}) = 1.0$ and find $a_i(\mathbf{t})$'s by (5.13b) that satisfy the origin constraints $\Sigma_j t_{ij} = o_i$, $i = 1, 2, \ldots I$;

(2) With the latest $a_i(\mathbf{t})$'s and by (5.13c), find $b_j(\mathbf{t})$'s which satisfy the destination constraints $\Sigma_i t_{ij} = d_j$, $j = 1, 2, \ldots J$;

(3) Keeping the $b_j(\mathbf{t})$'s fixed, calculate $a_i(\mathbf{t})$'s, again by (5.13b);

(4) Repeat steps (2) and (3) until convergence is achieved.

Once $a_i(\mathbf{t})$'s and $b_j(\mathbf{t})$'s are determined, the $F_{ij}(\mathbf{t})$'s can be obtained by (5.13a). Thus the numerical calculations of the doubly constrained model involve two nested iterations. The inner iteration is the one outlined above to obtain $a_i(\mathbf{t})$'s, $b_j(\mathbf{t})$'s so as to get $F_{ij}(\mathbf{t})$'s; the outer iteration is $\mathbf{t}(n+1) = \mathbf{F}(\mathbf{t}(n))$, made for n.

Numerical calculations have shown that the dynamic behaviour in the doubly constrained models is very similar to that in the unconstrained or singly constrained models. When the exponential and power deterrence functions are used, there are only point attractors and period-two attractors. When the combined deterrence function is used and when the dimension is lower, there are still only point and period-two attractors.

When the dimension is higher (4 or more) and when the combined deterrence function is used, the behaviour is more complicated. Chaos has been found to exist widely in the model. Experiments with a model of three origins and three destinations will be described here. The uncongested travel costs and capacities between the zones are as follows

$$\mathbf{c}^0 = \begin{bmatrix} c^0_{11} & c^0_{12} & c^0_{13} \\ c^0_{21} & c^0_{22} & c^0_{23} \\ c^0_{31} & c^0_{32} & c^0_{33} \end{bmatrix} = \begin{bmatrix} 1.00 & 1.00 & 1.50 \\ 1.20 & 1.40 & 1.80 \\ 1.50 & 0.90 & 0.50 \end{bmatrix},$$

$$\mathbf{q} = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} = \begin{bmatrix} 0.01 & 0.11 & 0.07 \\ 0.09 & 0.09 & 0.11 \\ 0.10 & 0.09 & 0.07 \end{bmatrix}.$$

The total numbers of trips from and to each zone are

$$\mathbf{o} = (o_1\ o_2\ o_3) = (0.35\ 0.35\ 0.30), \text{ and } \mathbf{d} = (d_1\ d_2\ d_3) = (0.30\ 0.30\ 0.40),$$

respectively. Figures 5.7a–5.7b show one of the chaotic attractors found in the model. The time series appears to be even more irregular than that for the

chaotic attractor in the unconstrained or singly constrained model (Figure 5.2e). The power spectrum is continuous, implying the motion is chaotic.

Bifurcation diagrams are produced for all the parameters in the model. These diagrams are even more time consuming to produce than those for the unconstrained or singly constrained model because each iteration of the equation involves a subiteration to solve the model.

Figure 5.8 is a bifurcation diagram for μ. It appears that there is no obvious periodic doubling sequence or any other clear bifurcation route in the diagram. Another feature of this graph is that there are a lot of broken points. This means that two or more attractors coexist in the phase space for the same values of parameters. Bifurcation diagrams are also made for all other parameters and they are similar to Figure 5.8.

Models with four origins and four destinations are also examined and similar behaviour found. It seems that the behaviour in the doubly constrained models is a little more complicated than that in the unconstrained or singly constrained models. This may because the phase space in the doubly constrained model is more complicated, as mentioned in the introduction of this chapter.

5.4. CALCULATION OF LIAPUNOV EXPONENTS

In this and the next section, chaotic behaviour found in the gravity model is examined further. As mentioned in the introduction of the thesis, chaos is a new kind of irregular behaviour found in deterministic systems. It is not yet fully understood. Therefore, it is of theoretical interest to find out if a model possesses chaos and to examine chaotic attractors.

Chaotic attractors are the attractors with *sensitive dependence on initial conditions* (Eckmann and Ruelle, 1985). A tiny difference in initial conditions can grow exponentially with time so that the state of the system is essentially unknown after a short time because of the uncertainties in the initial state. The divergence of neighbouring trajectories can be measured by Liapunov exponents. Liapunov exponents will be introduced first and then calculated for chaotic attractors of the gravity model. The theoretical basis of this section comes from Eckmann and Ruelle (1985).

5.4.1. Definition and algorithm

A Liapunov exponent is a measure of the average rate of change of small separations on an attractor. Consider the one dimensional discrete time evolution equation

$$x(n+1) = g(x(n)), \qquad x(i)\in\mathbb{R},$$

where n is the discrete time. The small initial separation $\delta x(0)$ after time N is then

$$\begin{aligned}\delta x(N) &= g^N(x(0)+\delta x(0)) - g^N(x(0)) \\ &\cong [\frac{\mathrm{d}}{\mathrm{d}x}(g^N)(x(0))]\ \delta x(0),\end{aligned}$$

where $g^N(x) = g(g(\cdots g(x)\cdots))$, N times. By the chain rule of differentiation, we have

$$\frac{\mathrm{d}}{\mathrm{d}x}(g^N)(x(0)) = \frac{\mathrm{d}}{\mathrm{d}x} g(x(N-1)) \times \frac{\mathrm{d}}{\mathrm{d}x} g(x(N-2)) \cdots \frac{\mathrm{d}}{\mathrm{d}x} g(x(0)).$$

Suppose that the separation grows (or decays) exponentially with N, that is,

$$\delta x(N) = \delta x(0)\ \mathrm{e}^{\lambda N} \cong [\frac{\mathrm{d}}{\mathrm{d}x}(g^N)(x(0))]\ \delta x(0).$$

Then the average rate of change or the Liapunov exponent is defined as (Eckmann and Ruelle, 1985)

$$\lambda = \lim_{N\to\infty} \frac{1}{N} \log \left|\frac{\mathrm{d}}{\mathrm{d}x}(g^N)(x(0))\right| = \lim_{N\to\infty} \frac{1}{N} \Sigma_{i=0}^{N-1} \log \left|\frac{\mathrm{d}}{\mathrm{d}x} g(x(i))\right|.$$

In the case of multi-dimensional systems, the derivative $(\mathrm{d}/\mathrm{d}x)g$ is replaced by the Jacobian matrix. The definition of Liapunov exponents for multi-dimensional systems will be explained using the gravity model $\mathbf{t}(n+1) = \mathbf{F}(\mathbf{t}(n))$.

Let $\mathbf{J}(\mathbf{t})$ be the Jacobian matrix of $\mathbf{F}(\mathbf{t})$ at the point $\mathbf{t} \in S$: $\mathbf{J}(\mathbf{t}) = \partial\mathbf{F}(\mathbf{t})/\partial\mathbf{t} \equiv [\partial F_i(\mathbf{t})/\partial t_j]$. Denote the Jacobian matrix $\partial\mathbf{F}^n(\mathbf{t})/\partial\mathbf{t}$ of the nth iteration $\mathbf{F}^n(\mathbf{t})$

by $\mathbf{J}_{\mathbf{t}}^{n}$. Then

$$\mathbf{J}_{\mathbf{t}}^{n} = \mathbf{J}(\mathbf{F}^{n-1}(\mathbf{t})) \ \ldots \ \mathbf{J}(\mathbf{F}(\mathbf{t})) \ \mathbf{J}(\mathbf{t}) \tag{5.14}$$

by the chain rule of differentiation. Here the superscript n denotes the number of iterations, not a power, and $\mathbf{J}(\mathbf{F}^{r}(\mathbf{t}))$ is the Jacobian matrix of $\mathbf{F}$ evaluated at the point $\mathbf{F}^{r}(\mathbf{t})$. Let σ_i be the ith eigenvalue of the matrix

$$\lim_{n\to\infty} [\mathbf{J}_{\mathbf{t}}^{n*} \ \mathbf{J}_{\mathbf{t}}^{n}]^{1/2n},$$

where $\mathbf{J}_{\mathbf{t}}^{n*}$ stands for the transpose of $\mathbf{J}_{\mathbf{t}}^{n}$. Then the ith Liapunov exponent is defined as (Eckmann and Ruelle, 1985)

$$\lambda_i = \text{Log} \ |\sigma_i|.$$

There are as many Liapunov exponents as the dimension of the phase space of a dynamical system. A positive exponent indicates expansion of the neighbouring trajectories and a negative exponent indicates contraction. The largest exponent is positive for chaotic attractors; it accounts for the sensitive dependence on initial conditions. A non-chaotic attractor does not possess positive exponents.

There are two methods for calculating Liapunov exponents, suggested by Wolf *et al.* (1985) and Eckmann and Ruelle (1985) respectively. The second algorithm is usually preferred (see, for example, Conte and Dubois, 1988) and is used here. The basis of the algorithm is to calculate the product (5.14) by QR factorizations (Hager, 1988). Let

$$\begin{aligned}
&\mathbf{Q}_1\mathbf{R}_1 = \text{the QR factorization of } \mathbf{J}(\mathbf{t}),\\
&\mathbf{Q}_2\mathbf{R}_2 = \text{the QR factorization of } \mathbf{J}(\mathbf{F}(\mathbf{t}))\ \mathbf{Q}_1,\\
&\ldots\ \ldots\\
&\mathbf{Q}_k\mathbf{R}_k = \text{the QR factorization of } \mathbf{J}(\mathbf{F}^{k-1}(\mathbf{t}))\ \mathbf{Q}_{k-1},\\
&\ldots\ \ldots\ ,
\end{aligned}$$

where $\mathbf{Q}_k$ is an orthogonal matrix and $\mathbf{R}_k$ an upper triangular matrix with non-negative diagonal elements. Then

$$\mathbf{J}(\mathbf{F}^{n-1}(\mathbf{t})) \ \ldots \ \mathbf{J}(\mathbf{F}(\mathbf{t}))\ \mathbf{J}(\mathbf{t}) = \mathbf{Q}_n \mathbf{R}_n \ \ldots \ \mathbf{R}_1$$

The diagonal elements $(\nu_n)_{ii}$ of the upper triangular matrix product $\mathbf{R}_n \ldots \mathbf{R}_1$ lead to the exponents (Eckmann and Ruelle, 1985)

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \text{Log} \, |\text{diagonal elements } (\nu_n)_{ii}|,$$

where n is the number of iterations.

To implement the algorithm, the gravity model is iterated. At each iteration, the Jacobian matrix is calculated and is multiplied by the previous orthogonal matrix. The product is then decomposed by the QR factorization to get the upper triangular matrix, from which the eigenvalues are obtained. The iterations are continued until a convergence to the Liapunov exponents is achieved. The algorithm was programmed in FORTRAN, making use of NAG routines (The NAG Ltd., 1987). See Appendix B for the program listing.

The program was tested by calculating the Liapunov exponents of the chaotic attractor of the Hénon map shown in Figure 3.3 in Chapter 3. The Liapunov exponents for this attractor are known to be $\lambda_1 = 0.42$, $\lambda_2 = -1.6$ (Conte and Dubois, 1988). The first 2,000 steps of calculation of the Liapunov exponents for this attractor are shown in Figure 5.9, where it can be seen that they converge quickly. The result after 20,000 iterations is [0.4168 –1.6208].

5.4.2. Calculation for the gravity models

The algorithm described in the last subsection can be used directly to calculate the Liapunov exponents for the unconstrained or the singly constrained gravity model. Liapunov exponents for the chaotic attractor shown in Figure 5.2 were calculated. The first 5,000 iterations are shown in Figure 5.10; the convergence is apparent. The three exponents after 30,000 iterations are [0.20 –0.02 –0.70], with the first one being positive. Shown in Figures 5.11 is the first Liapunov exponents calculated as a function of β, in company with the bifurcation diagram in Figure 5.5a. The second and the third exponents are all negative. By comparing Figure 5.11 with Figure 5.5a it can be seen that the first exponents are negative for non-chaotic attractors and are positive for chaotic ones.

For the doubly constrained gravity model, the calculation of Liapunov exponents needs a little more treatment. Because the two sets of normalizing factors in the doubly constrained model are interdependent, the partial derivatives in the Jacobian matrix are not directly available. These, however, can be found indirectly, as we shall see now.

Consider the doubly constrained model (5.13)

$$F_{ij}(\mathbf{t}) = a_i(\mathbf{t}) b_j(\mathbf{t}) f(c_{ij}(t_{ij})), \quad t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i, \ \Sigma_i t_{ij} = d_j,$$

where $a_i(\mathbf{t})$ and $b_j(\mathbf{t})$ satisfy the equations

$$a_i(\mathbf{t}) = o_i \frac{1}{\Sigma_j b_j(\mathbf{t}) f(c_{ij}(t_{ij}))}, \quad i = 1, 2, \ldots I,$$

and

$$b_j(\mathbf{t}) = d_j \frac{1}{\Sigma_i a_i(\mathbf{t}) f(c_{ij}(t_{ij}))}, \quad j = 1, 2, \ldots J.$$

The partial derivatives can be found to be:

$$\frac{\partial F_{ij}}{\partial t_{kl}} = b_j(\mathbf{t}) f(c_{ij}(t_{ij})) \frac{\partial}{\partial t_{kl}} a_i(\mathbf{t}) + a_i(\mathbf{t}) f(c_{ij}(t_{ij})) \frac{\partial}{\partial t_{kl}} b_j(\mathbf{t})$$

$$+ a_i(\mathbf{t}) b_j(\mathbf{t}) \frac{\partial}{\partial t_{kl}} f(c_{ij}(t_{ij})), \tag{5.15}$$

where

$$\frac{\partial}{\partial t_{kl}} f(c_{ij}(t_{ij})) = \begin{cases} \frac{\mathrm{d}}{\mathrm{d}t_{kl}} f(c_{ij}(t_{ij})), & \text{if} \quad i=k, \ j=l \\ 0 & \text{otherwise} \end{cases}.$$

By (5.13a) and (5.13b), $\frac{\partial}{\partial t_{kl}} a_i(\mathbf{t})$ and $\frac{\partial}{\partial t_{kl}} b_j(\mathbf{t})$ should satisfy

$$\frac{\partial}{\partial t_{kl}} a_i(\mathbf{t}) = -\frac{[a_i(\mathbf{t})]^2}{o_i} \Sigma_j f(c_{ij}(t_{ij})) \frac{\partial}{\partial t_{kl}} b_j(\mathbf{t}) - \frac{[a_i(\mathbf{t})]^2}{o_i} \Sigma_j b_j(\mathbf{t}) \frac{\partial}{\partial t_{kl}} f(c_{ij}(t_{ij}))$$

and

$$\frac{\partial}{\partial t_{kl}} b_j(\mathbf{t}) = -\frac{[b_j(\mathbf{t})]^2}{d_j} \Sigma_i f(c_{ij}(t_{ij})) \frac{\partial}{\partial t_{kl}} a_i(\mathbf{t}) - \frac{[b_j(\mathbf{t})]^2}{d_j} \Sigma_i a_i(\mathbf{t}) \frac{\partial}{\partial t_{kl}} f(c_{ij}(t_{ij}))$$

for $i, k = 1, 2, ..., I,$ and $j, l = 1, 2, ..., J.$

There are $(I+J)\times I\times J$ equations which are linear with respect to the same number of unknowns, $\frac{\partial}{\partial t_{kl}} a_i(\mathbf{t})$, $\frac{\partial}{\partial t_{kl}} b_j(\mathbf{t})$. So they can be solved numerically to obtain $\frac{\partial}{\partial t_{kl}} a_i(\mathbf{t})$, $\frac{\partial}{\partial t_{kl}} b_j(\mathbf{t})$. The partial derivatives can then be found by (5.15) to get the Jacobian matrix. The above algorithm can now be used to calculate the Liapunov exponents. Liapunov exponents are calculated for the chaotic attractor shown in Figure 5.7. The first 5,000 iterations are shown in Figure 5.12. The final result after 20,000 iterations is [0.1248 –0.1449 –0.3597 –0.9477]. The first one is positive, which confirms the attractor is chaotic.

Sensitivity to initial conditions is a fundamental character of a chaotic attractor. Trajectories on the attractor expand in the directions with positive Liapunov exponents and contract in other directions with negative exponents. The directions of expansion and contraction are different at different point on the attractor. As a result, two closely spaced points are soon found very far apart. The diverging trajectories will fold back at certain point since the attractor is bounded. Therefore, a chaotic attractor is very complicated and normally possesses a fractal dimension. Fractal dimension is considered next.

5.5. CALCULATION OF FRACTAL DIMENSIONS

5.5.1. Definition and algorithm

Fractal dimensions are used to characterize the geometric feature of chaotic attractors. The dimension of an attractor is a lower bound to the number of state variables needed to describe a steady-state behaviour. It can quantify the complexity of an attractor. A strange attractor normally possesses non-integer dimension, called a *fractal dimension.* Several types of fractal dimension have been defined in the literature (Parker and Chua, 1989). However, it should be mentioned that fractal dimension is still an active research area, and the relationships and the meanings of different definitions of dimension are unclear,

especially in experimental settings and when applied to simulations. The main use of dimension here is to characterize the chaotic attractors by giving the minimum number of variables needed to describe the dynamics on the attractor. Bearing these facts in the mind, and considering that different definitions of dimension give values close to one another, there seems no theoretical reason for choosing one type of dimension over another. The main reason for choosing one type of dimension over another is the ease and accuracy of its computation. A commonly used algorithm for computing fractal dimension is the box-counting algorithm. However, it is very inefficient. For a system of dimension greater than three, the memory requirements are extensive and a large amount of data is needed. The algorithm for computing the *correlation dimension* due to Grassberger and Procaccia (1983) is more efficient and is therefore employed here to estimate correlation dimension for chaotic attractors in the gravity models.

The correlation dimension is defined based on the *correlation function* of an attractor. A correlation function is the average fraction of points within a certain radius r on the attractor. Let the sequence of N points

$$\{\mathbf{t}(1), ..., \mathbf{t}(n), ..., \mathbf{t}(N)\}$$

be an orbit on an attractor in the system (5.1). Then the correlation function $C(r)$ of the attractor is given by (Parker and Chua, 1989)

$$C(r) = \lim_{N\to\infty} \frac{1}{N^2} \{ \text{the number of points } (\mathbf{t}(i), \mathbf{t}(j)) \text{ such that } |\mathbf{t}(i) - \mathbf{t}(j)| < r \}.$$

The correlation dimension D_C is defined as (Grassberger and Procaccia, 1983)

$$D_C = \lim_{r\to 0} \frac{\log\ C(r)}{\log\ r} .$$

This definition may be interpreted geometrically. If r is the diameter of a volume element in the attractor, then the correlation function, $C(r)$, is proportional to r^{D_C}, and the above definition follows immediately.

To calculate the dimension, the model is first iterated and the transient removed to get an orbit on the attractor. Then the correlation functions are calculated for different values of r. For a given r, estimating $C(r)$ involves computing all

the inter-point distances, $|\mathbf{t}(i) - \mathbf{t}(j)|$, counting the number of points which are within the distance r for all points on the attractor, and dividing this number by N^2 to get $C(r)$. The correlation dimension is the slope of the plot of $\log(C(r))$ versus $\log(r)$. There may be only a limited range of the graph which is straight with an approximately constant slope. Only this range of data is used to estimate the dimension. The log-log plot of $C(r)$ versus r was made by the MATLAB software (The MathWorks Inc., 1993) and the slope is estimated by least squares, also in MATLAB. The algorithm requires the calculation of inter-point distances. Given N points, there are about $N^2/2$ distances that need to be computed. Since $N = 10{,}000$ is not uncommon, the distance calculations are the most time consuming part of the algorithm.

The program for this algorithm is tested by calculating the correlation dimension of the chaotic attractor in the Hénon map mentioned in Chapter 3. Figure 5.13 is the log-log plot of the correlation function versus the distances from the calculation. The slope or the correlation dimension estimated is 1.256, which agrees well to that in the literature, 1.261 (Grassberger and Procaccia, 1983).

5.5.2. Calculation for the gravity models

Consider the dimension for the chaotic attractor (Figure 5.2) found in the unconstrained or singly constrained gravity model. The log-log plot is shown in Figure 5.14 and the dimension estimated is 1.8251, though the attractor lies in a three dimensional phase space.

Also calculated is the dimension for the chaotic attractor shown in Figure 5.7 for the doubly constrained gravity model. The phase space of this model is four dimensional. Figure 5.15 shows the correlation function of the attractor. The slope, or the correlation dimension, is found to be 1.653.

The dynamic gravity model is a *dissipative system*: the phase volume shrinks with time. The phase space of the models considered above is three or four dimensional, but the evolution of the system is such that the final variations are confined within a region with a dimension of 1.7 or 1.8.

5.6. SUMMARY AND COMMENTS

In this chapter, the dynamics of the gravity model have been investigated. The existence, and the conditions for the uniqueness and the stability of the equilibria in the model were established theoretically. The equilibria in the models with the power or the exponential deterrence function were found to be unique. The unique equilibrium in the model with the exponential deterrence function gives the stochastic user equilibrium (SUE) for an O–D network. With the combined deterrence function, the equilibrium is unique if the condition (5.5) is satisfied.

Numerical calculations were made to find other types of dynamic behaviour. Point attractors and period-two attractors have been found to be the main feature in the models of lower dimensions with any of the three types of deterrence function, and in the models of higher dimensions with the power or the exponential deterrence function. When the dimension is higher (3 or more for unconstrained or singly constrained model, and 4 or more for the doubly constrained model), and when the combined deterrence functions is used, period doubling, chaos, and other complicated bifurcations were found in the gravity model. Liapunov exponents and correlation dimensions were calculated for the chaotic attractors and were found to be positive and fractal respectively.

The steady-state behaviour of the model depends on the inter dependence of flows and costs. The mechanism is that trip makers make their choices of origins and destinations based on travel costs, while the travel costs change with flows. If the values of the parameters are such that the interactions between flows and costs are mild enough, then an equilibrium can be reached. When the interactions become stronger, the behaviour of models becomes more complicated; oscillations and even chaos may occur.

It needs to be stressed that chaotic behaviour has been found in the gravity model with the combined deterrence function, which, according to (Ortúzar and Willumsen, 1990), can fit the observed data better than the other two forms of function. The difference between this function and the other two is that it is not monotonic while the other two are. This might be the reason why the combined function causes complicated behaviour in the gravity model. A drawback of the combined function, however, is that it has two parameters for calibration (determination of parameter values so that the model can fit the observed cost distribution as well as possible); the other two forms of deterrence function have

only one parameter to determine.

In practice, the values of parameters in the dynamic gravity model may be very different for different geographical areas. For a particular area, empirical studies are needed , preferably with the considerations of dynamic trip distributions, to determine the type of model and the values of parameters in the model. Then we could find out if the equilibrium would be stable and if the behaviour would be chaotic. In addition, these studies are also useful to find out whether the assumed dynamic model is realistic or not.

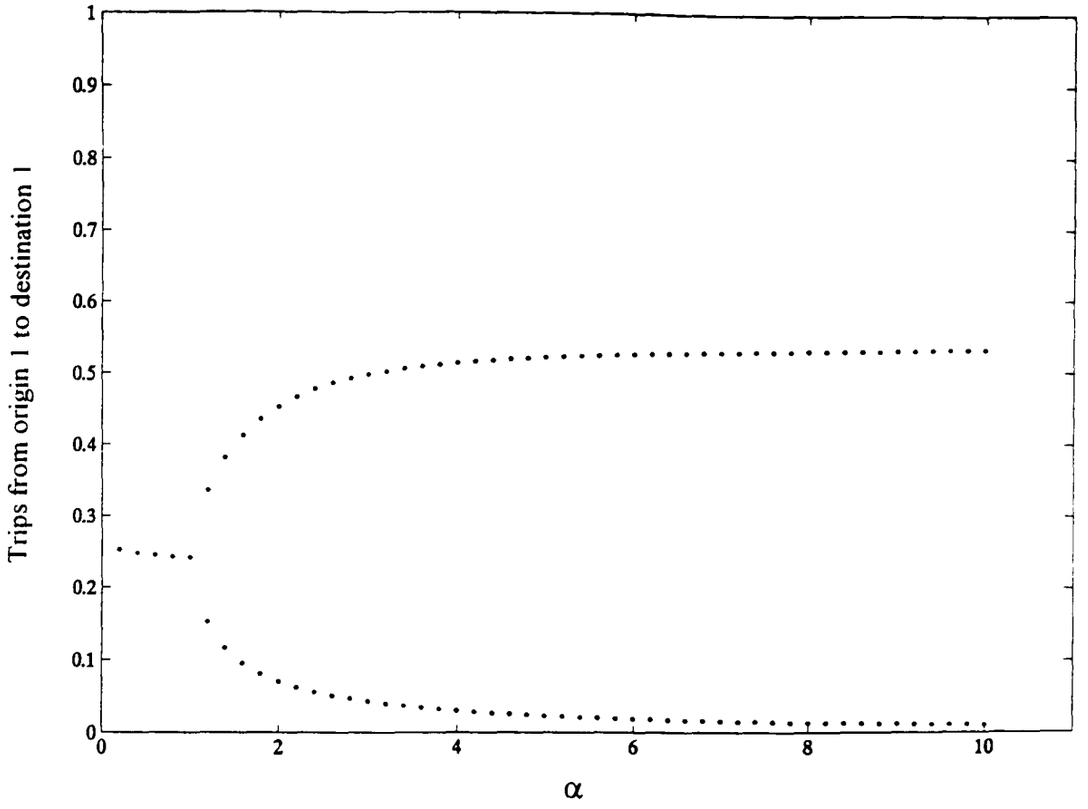
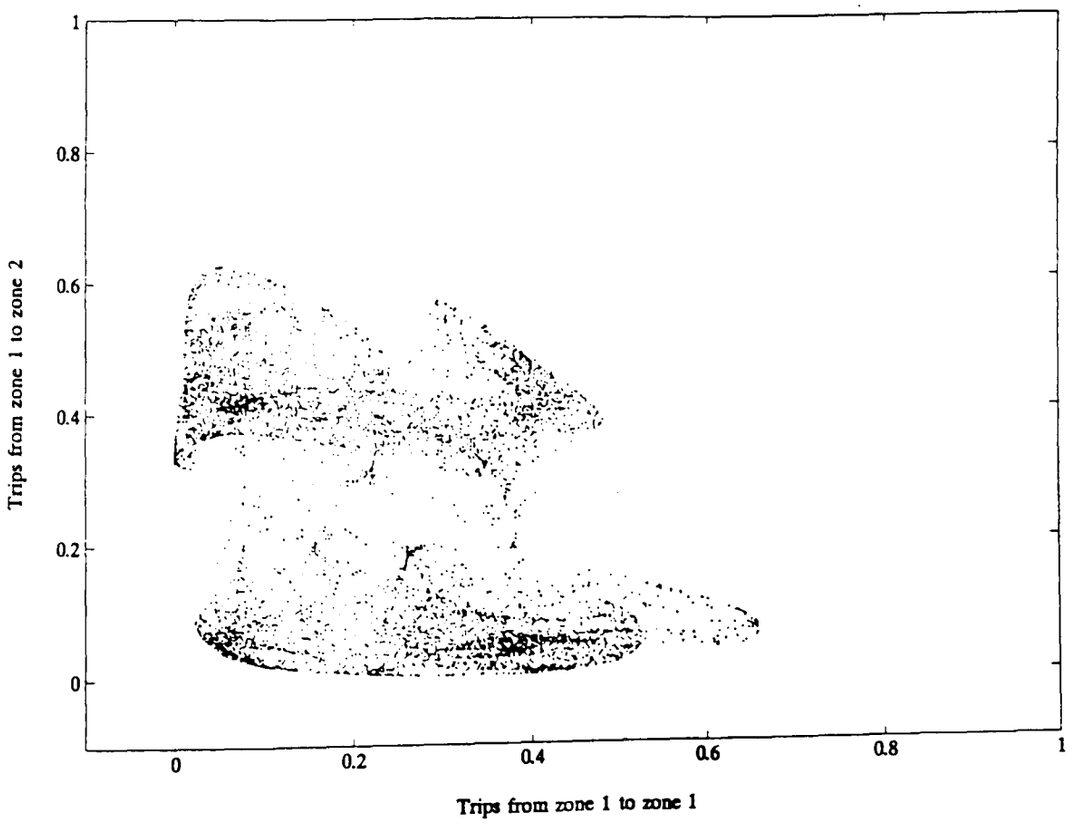
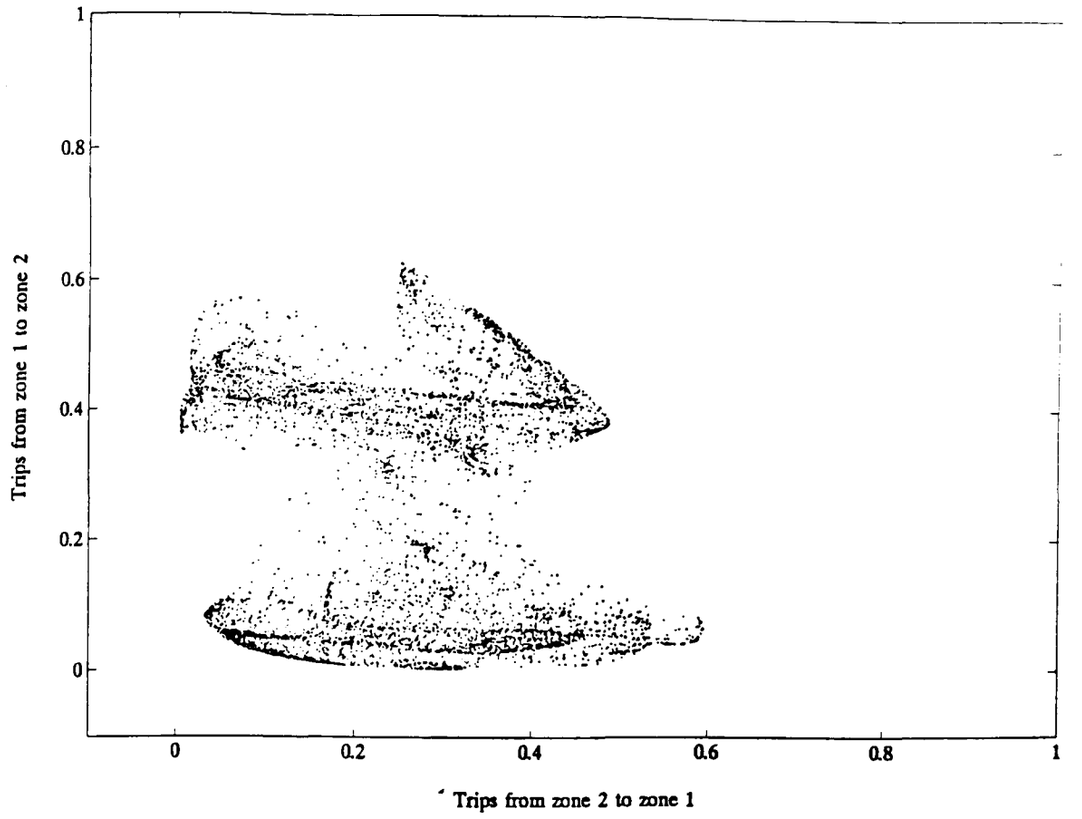


Figure 5.1 Bifurcation diagram of the unconstrained or singly constrained gravity model for α, with $\beta = 0$, $\mu = -1.0$, $\gamma = 1.5$.

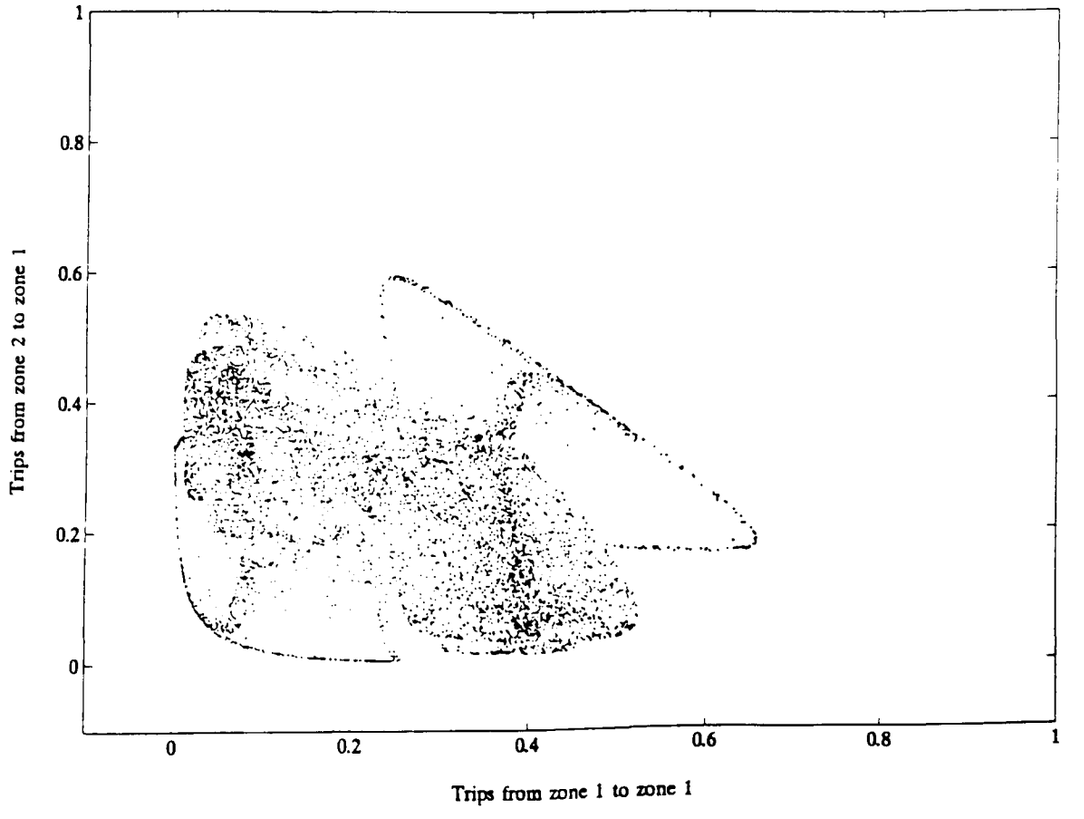


(a)

Figure 5.2 Chaotic attractor of the unconstrained or singly constrained gravity model, with $\mu = 8.0$, $\beta = 3.25$, $\alpha = 1.0$, $\gamma = 1.0$. (a) Phase portrait projection.

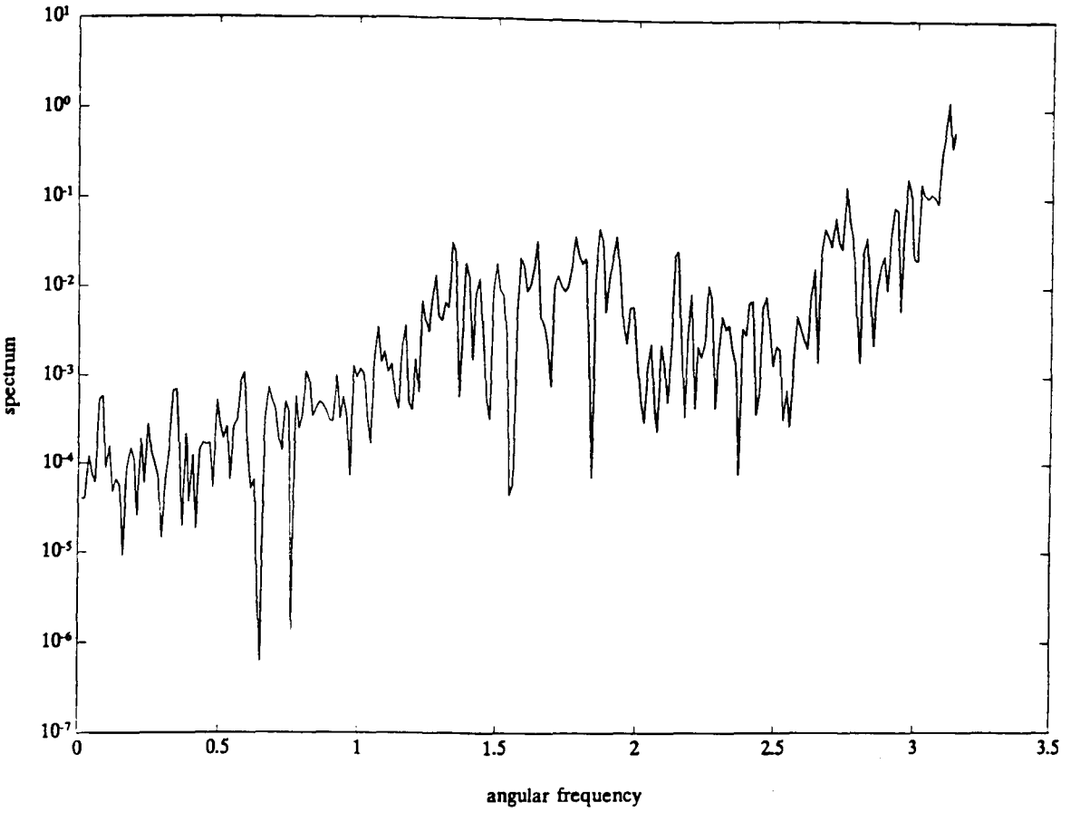


(b)

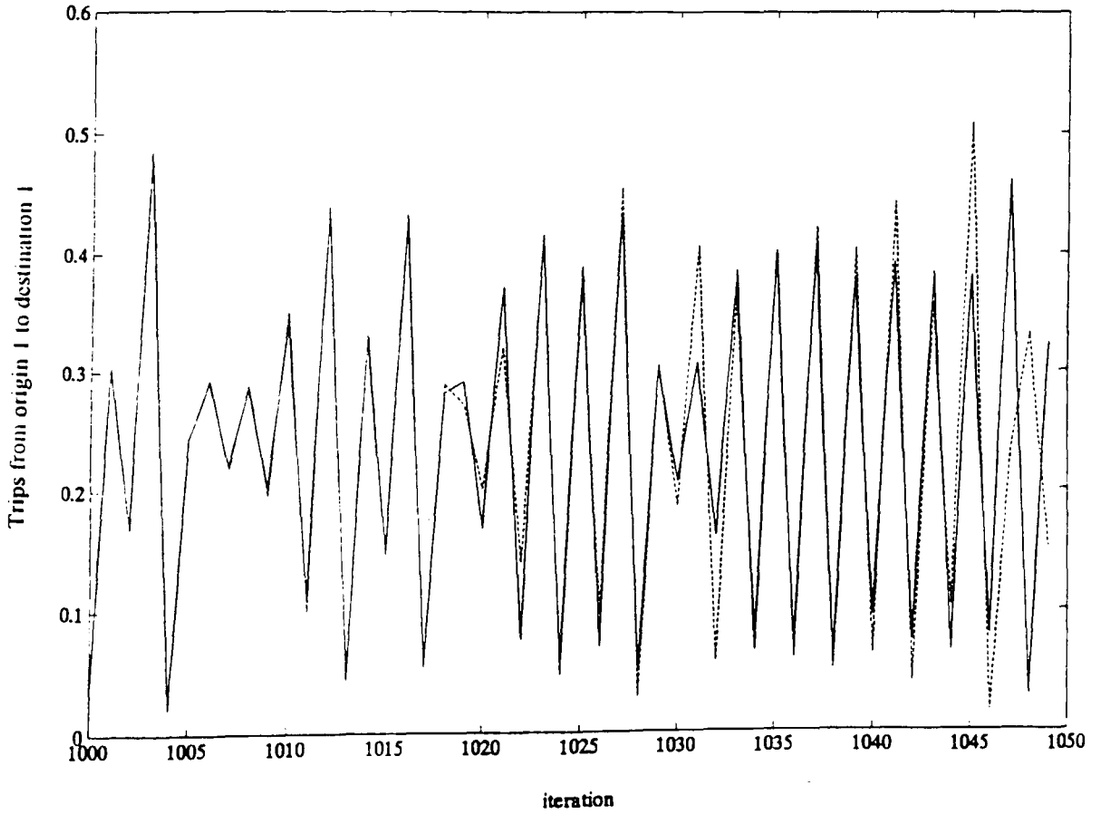


(c)

Figure 5.2 continued. (b)–(c) Phase portrait projections.

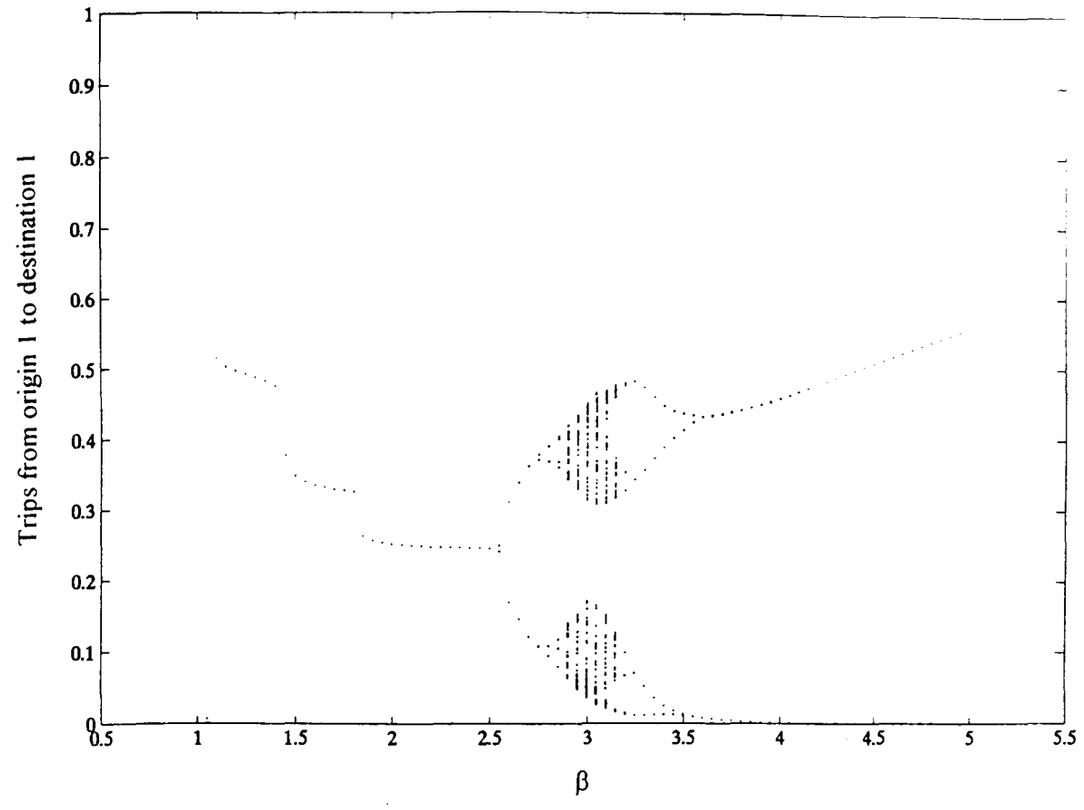


(d)

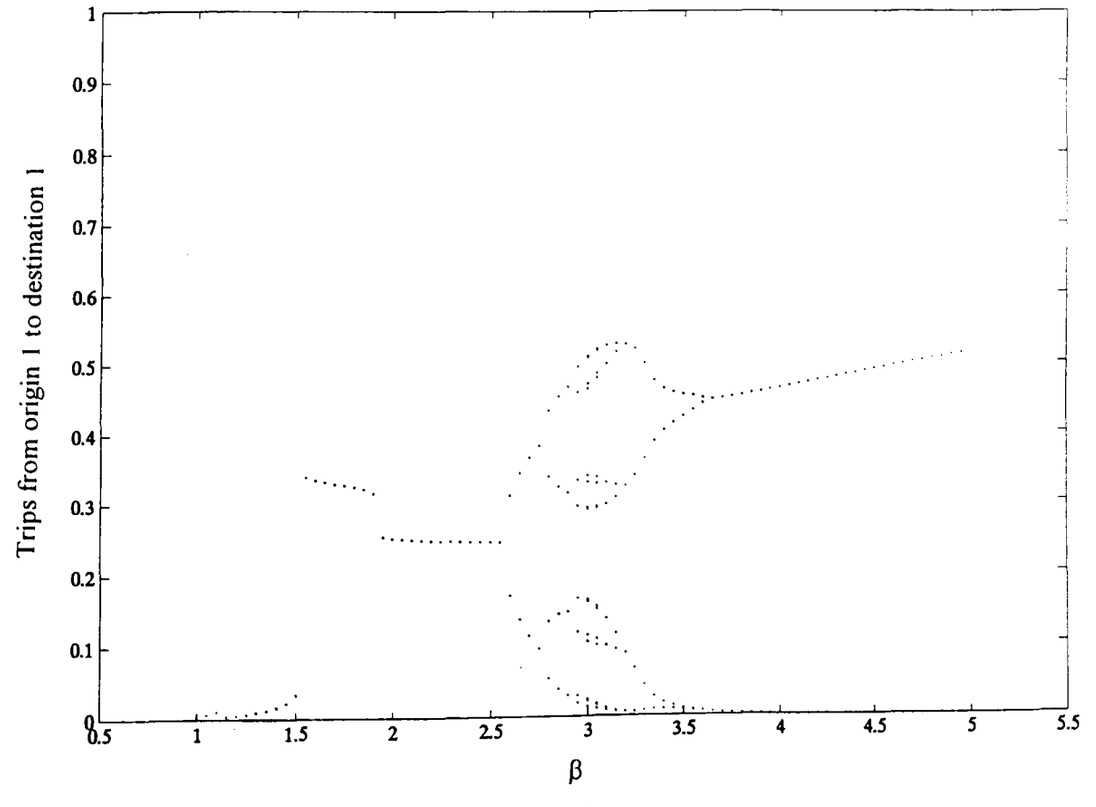


(e)

Figure 5.2 continued. (d) Power spectrum for the number of trips from origin 1 to destination 1; (e) sensitive dependence on initial conditions.

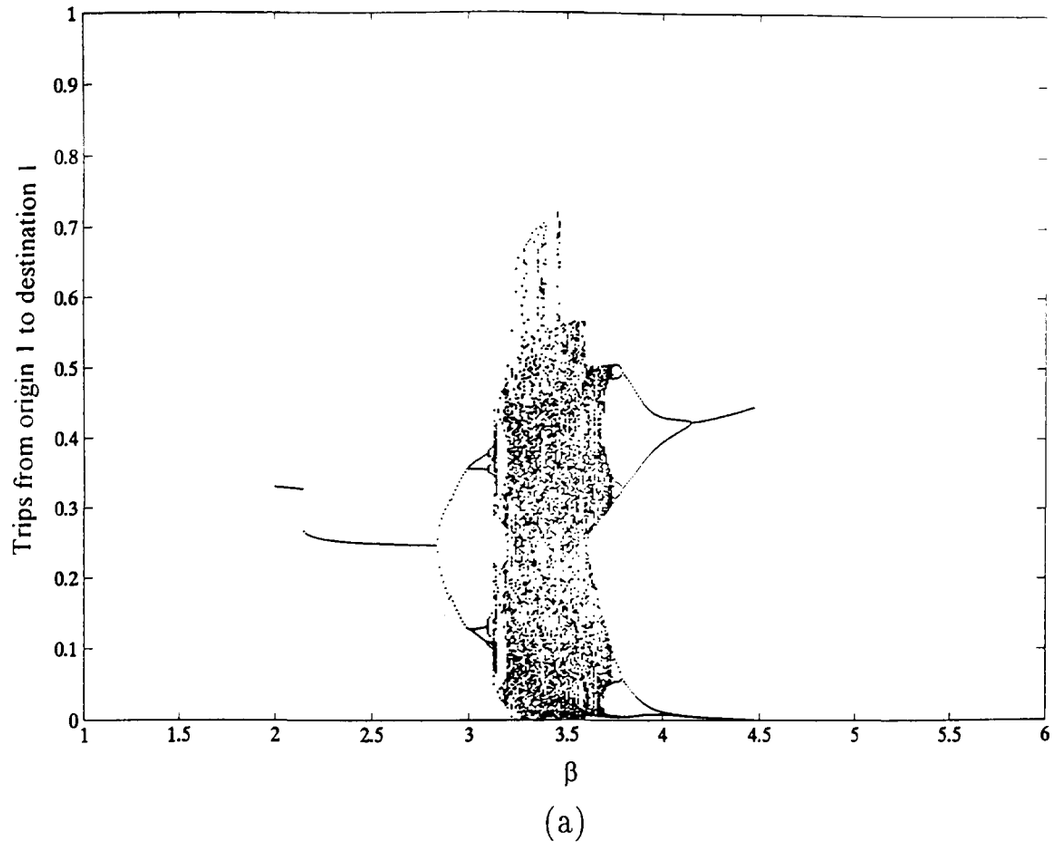


(a)



(b)

Figure 5.3 Bifurcation diagram of the unconstrained or singly constrained gravity model for β, with $\mu = 7.0$, $\alpha = 1.0$, $\gamma = 1.0$. (a) Initial conditions are the same for all values of β; (b) initial conditions are the final states of the previous step of β.



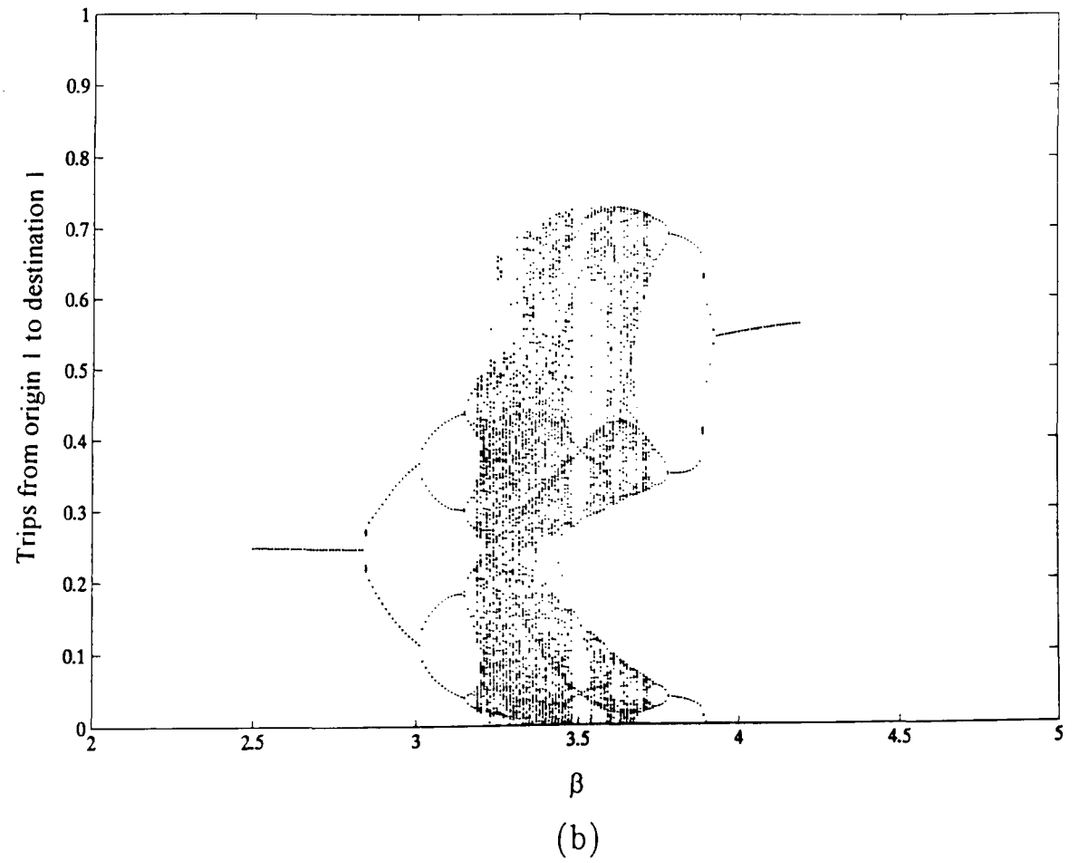
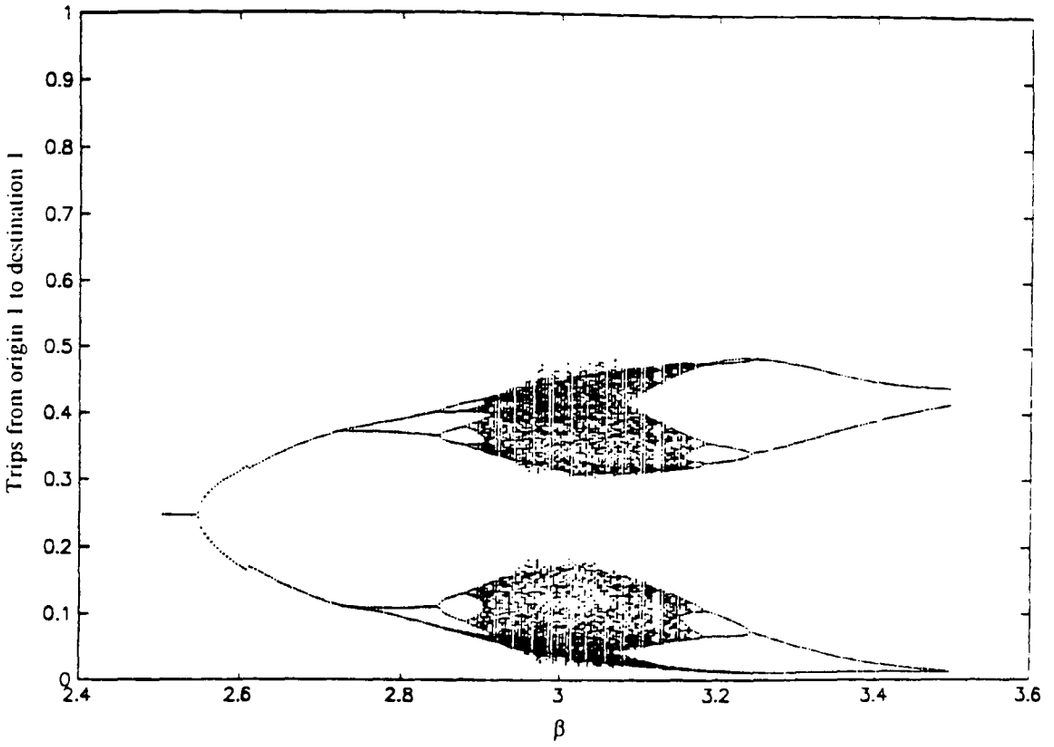
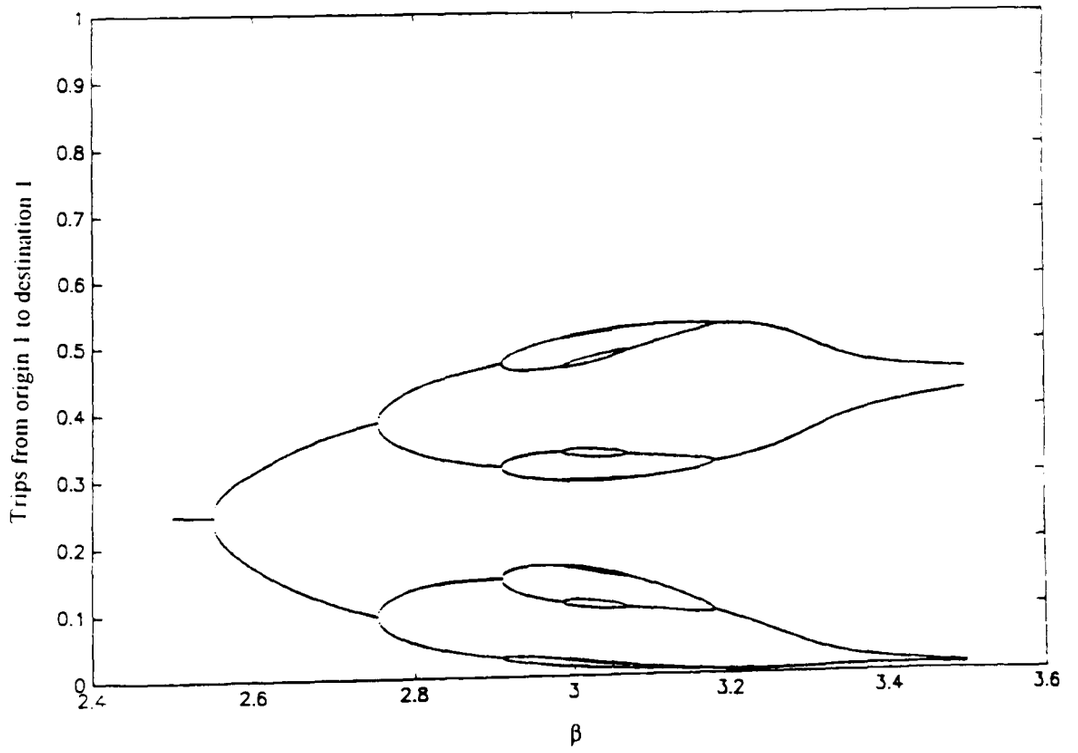


Figure 5.4 Bifurcation diagram of the unconstrained or singly constrained gravity model for β, with $\mu = 8.0$, $\alpha = 1.0$, $\gamma = 1.0$. (a) Initial conditions are the same for all values of β; (b) initial conditions are the final states of the previous step of β.

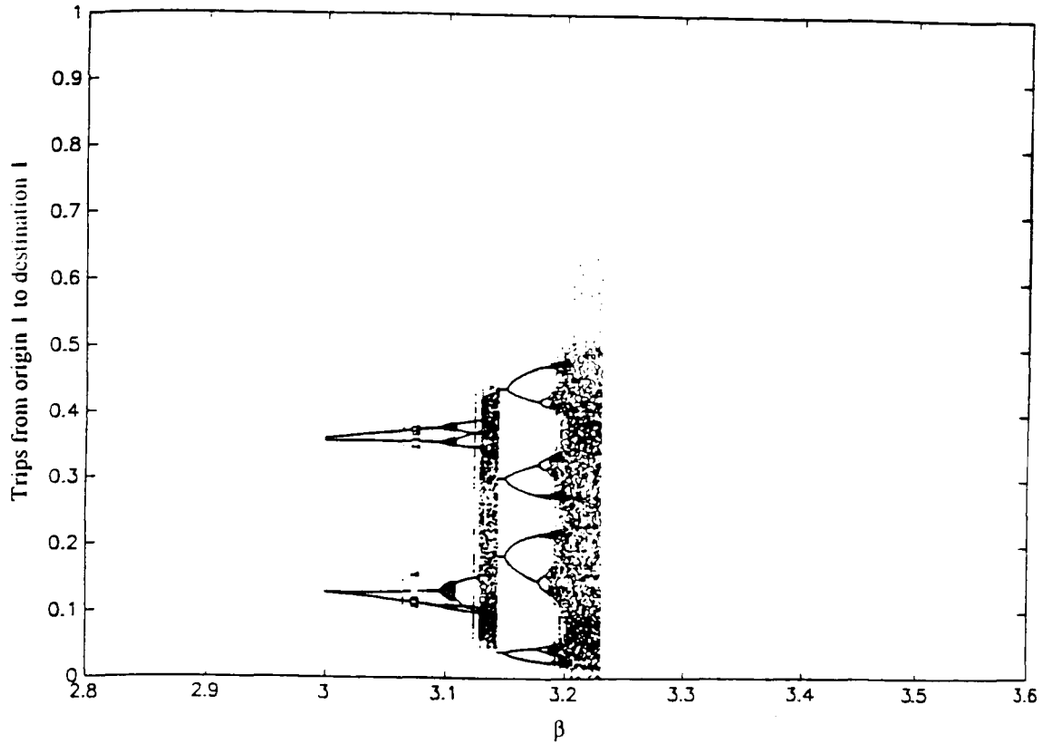


(a)

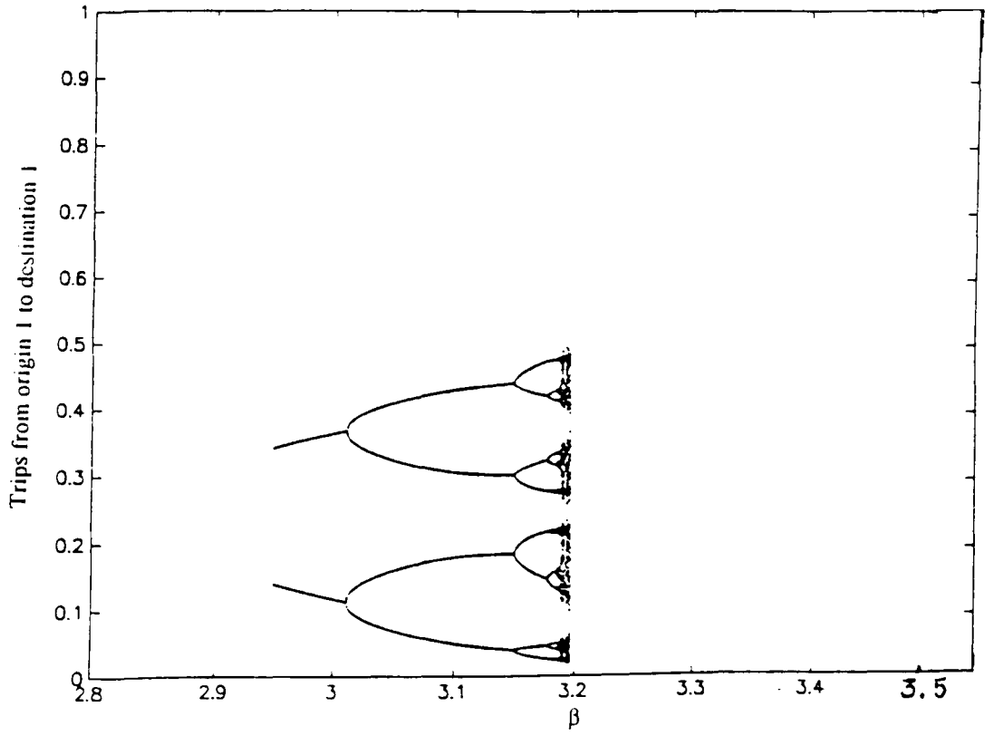


(b)

Figure 5.5 Enlargements of Figures 5.3a–5.3b. (a) Enlargement of Figure 5.3a; (b) enlargement of Figure 5.3b.

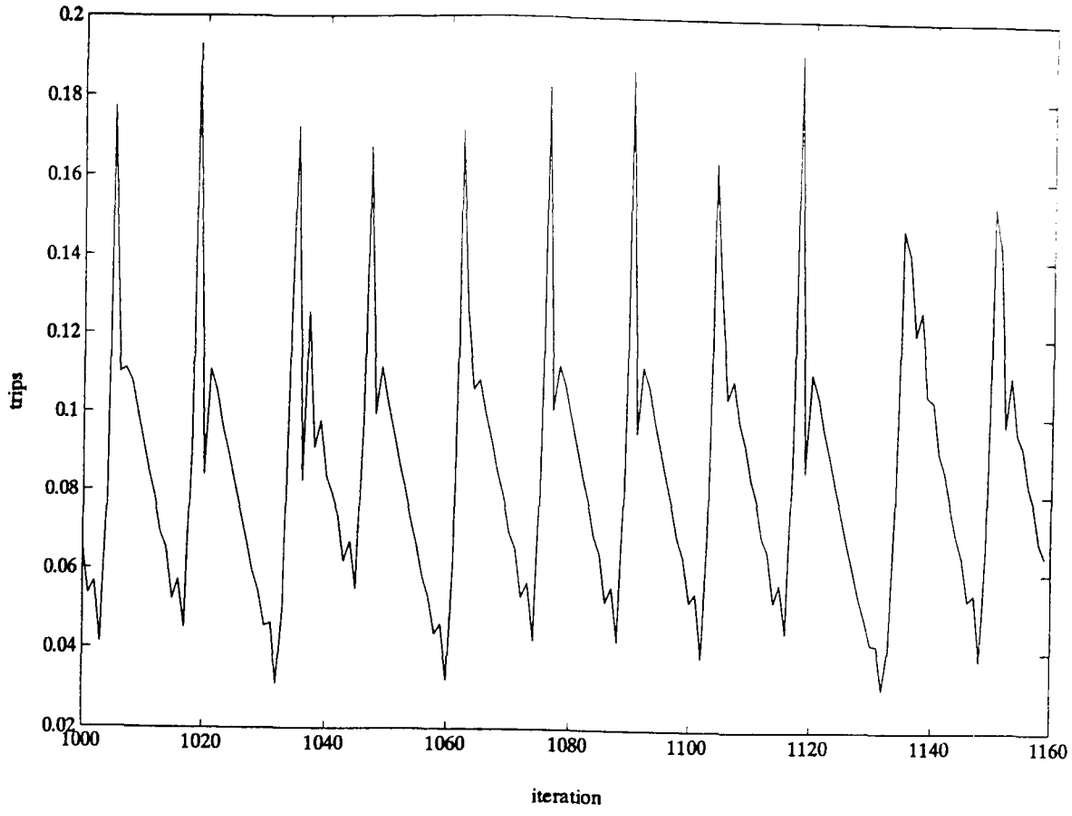


(a)

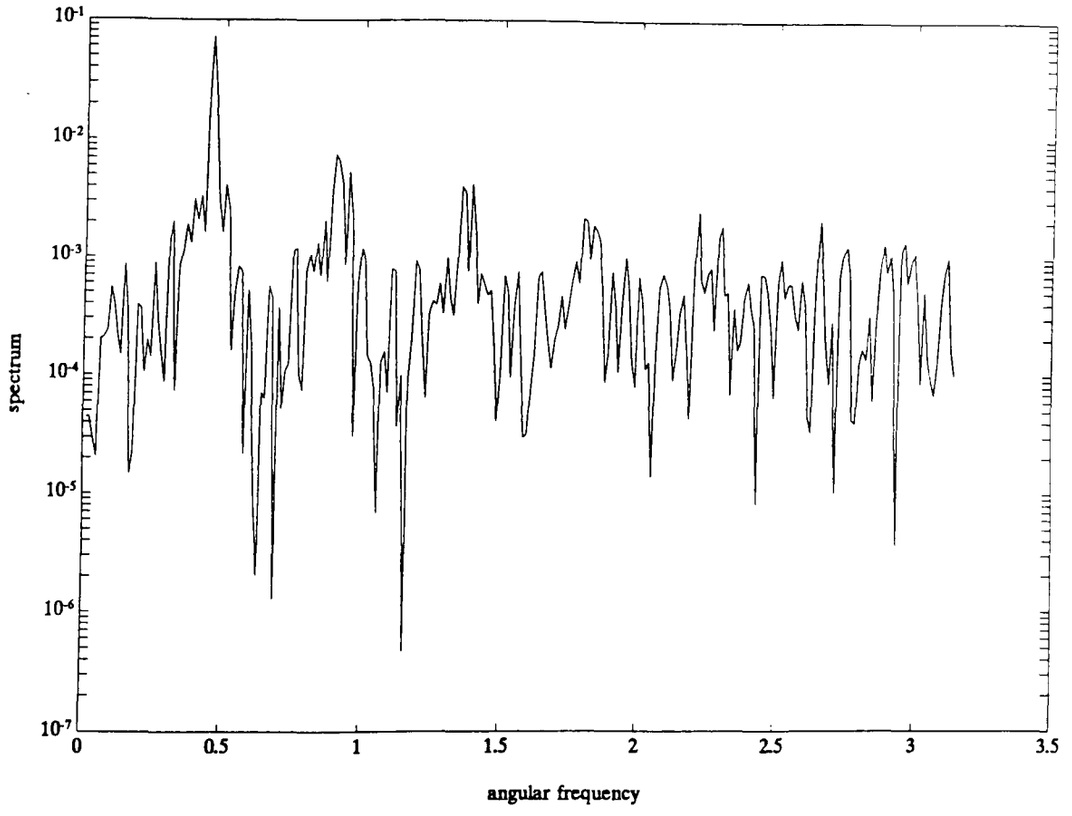


(b)

Figure 5.6 Enlargements of Figures 5.4a–5.4b. (a) Enlargement of Figure 5.4a; (b) enlargement of Figure 5.4b.



(a)



(b)

Figure 5.7 Chaotic attractor of the doubly constrained gravity model, with $\mu = 4.5$, $\beta = 1.25$, $\alpha = 1.5$, $\gamma = 1.5$. (a) Number of trips from origin 2 to destination 1; (b) power spectrum.

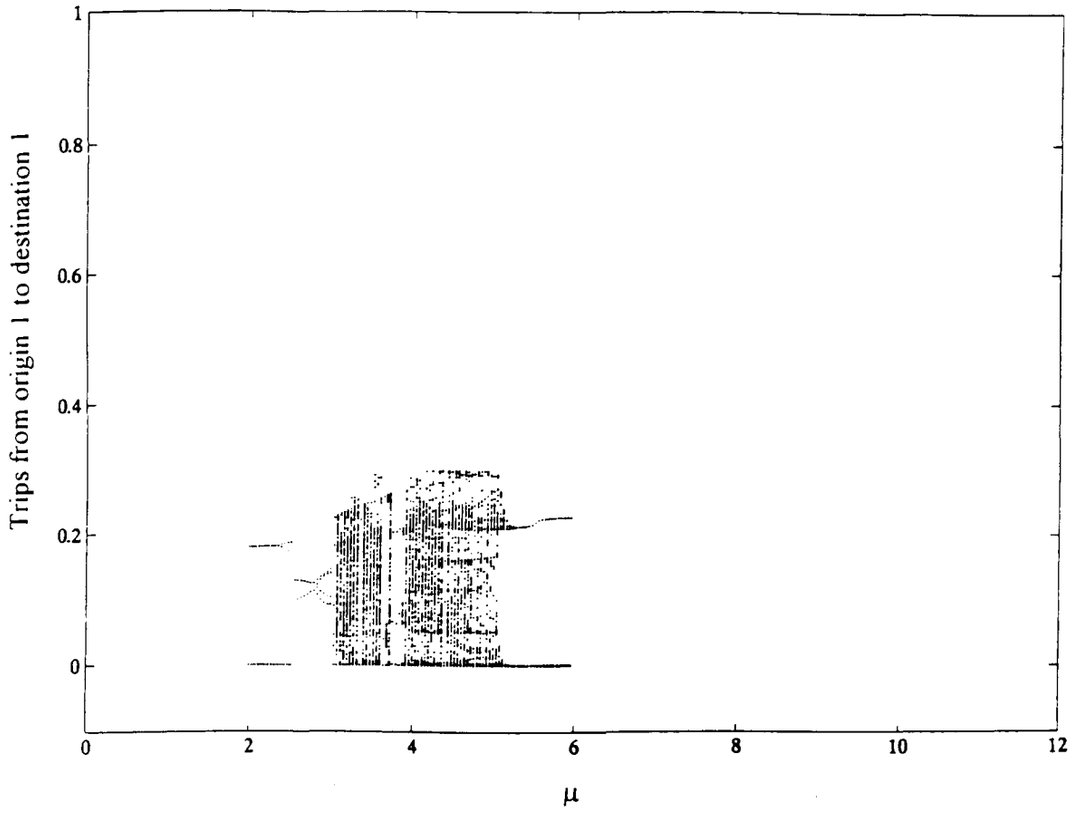


Figure 5.8 Bifurcation diagram of the doubly constrained gravity model for μ, with $\beta = 1.0$, $\alpha = 2.5$, $\gamma = 2.5$.

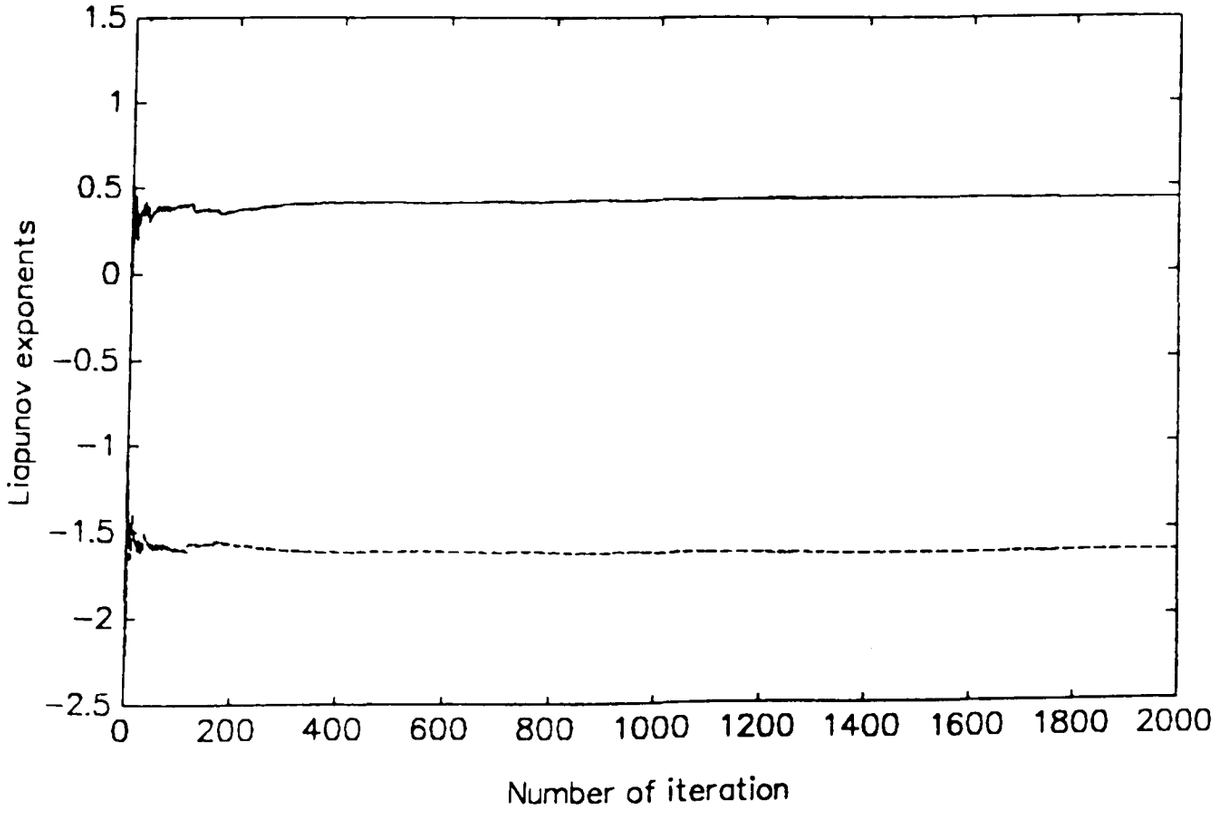


Figure 5.9 Convergence of Liapunov exponents for the chaotic attractor of the Hénon map.

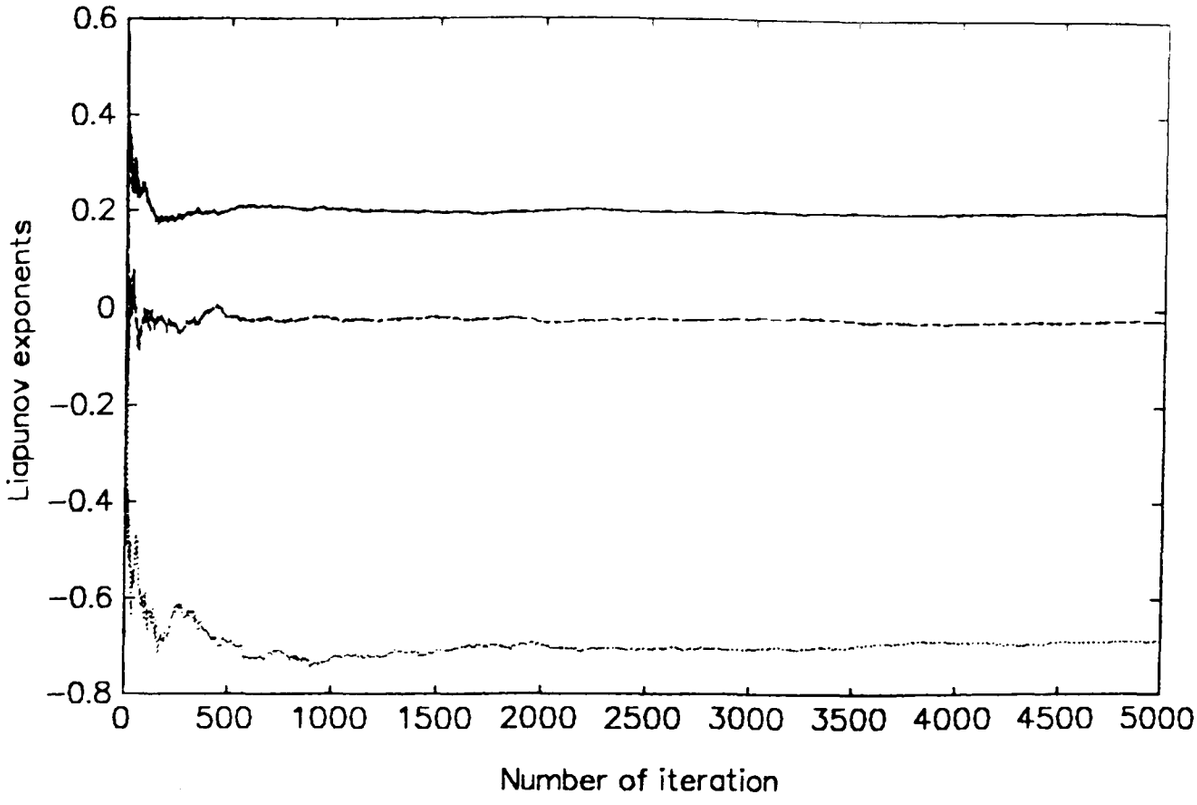


Figure 5.10 Convergence of Liapunov exponents for the chaotic attractor of the unconstrained or singly constrained gravity model.

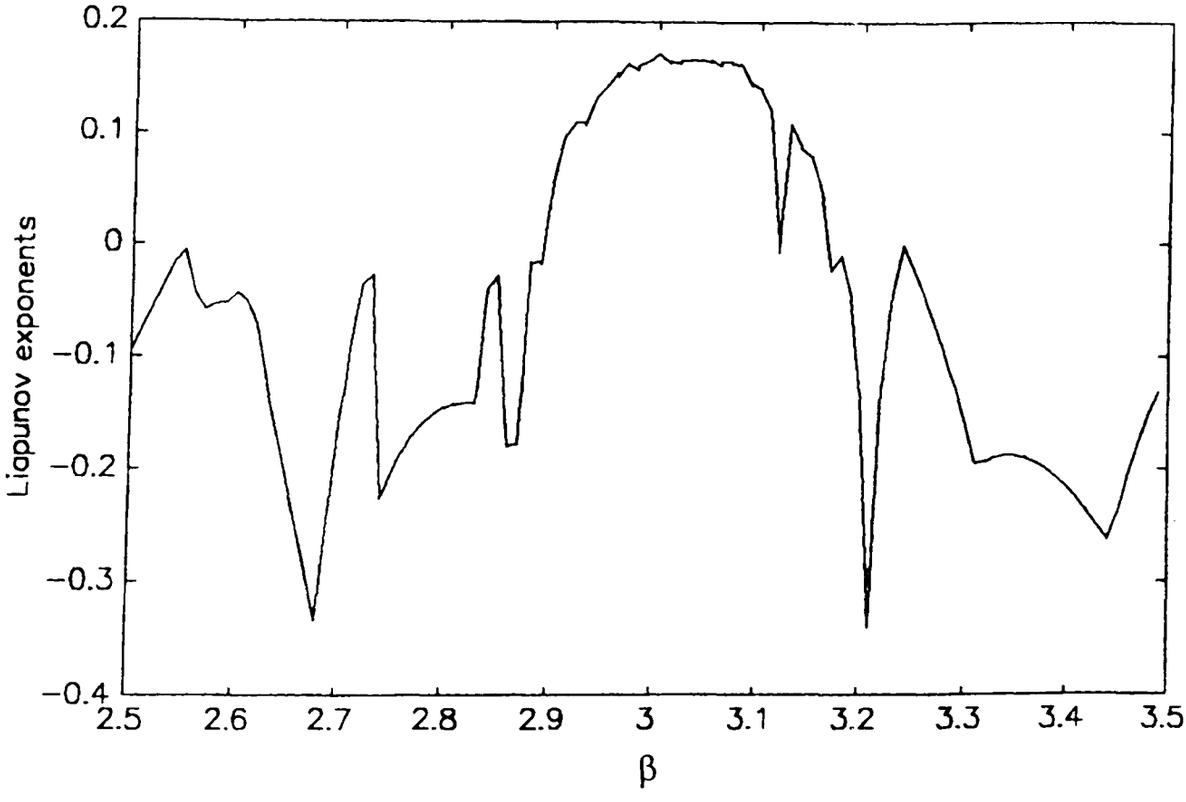


Figure 5.11 The first Liapunov exponents against β for the unconstrained or singly constrained gravity model.

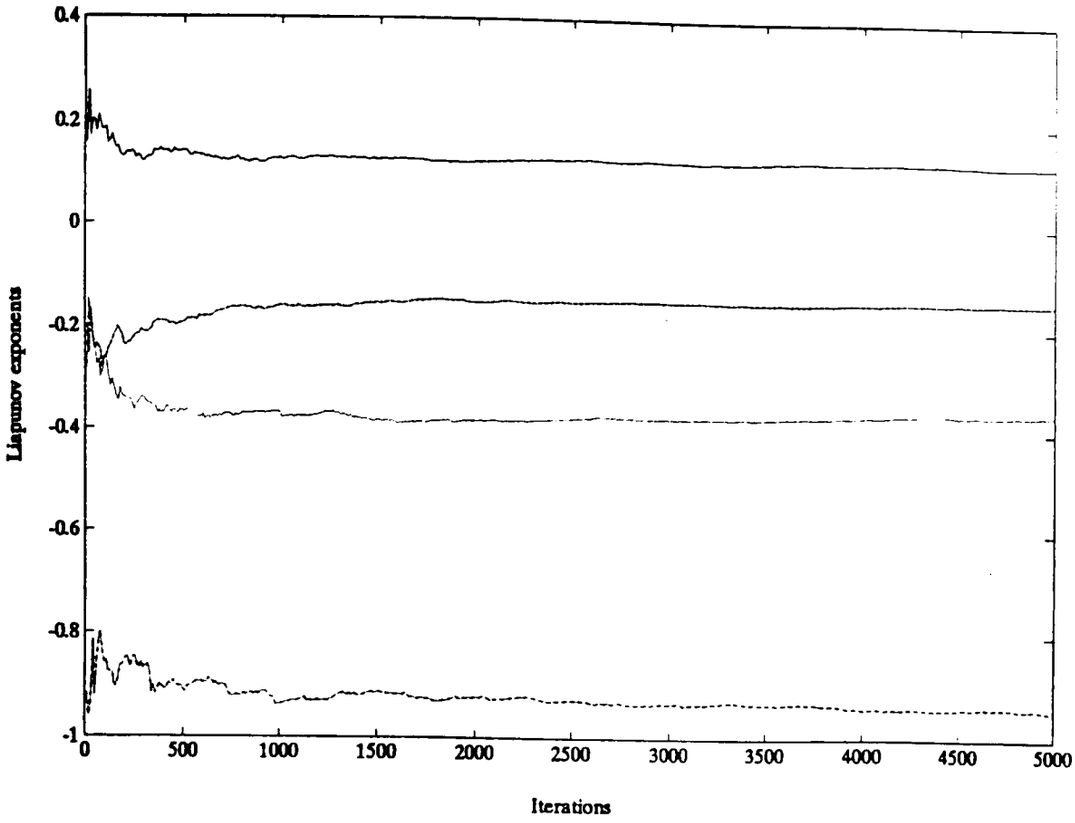


Figure 5.12 Convergence of Liapunov exponents for the chaotic attractor of the doubly constrained gravity model.

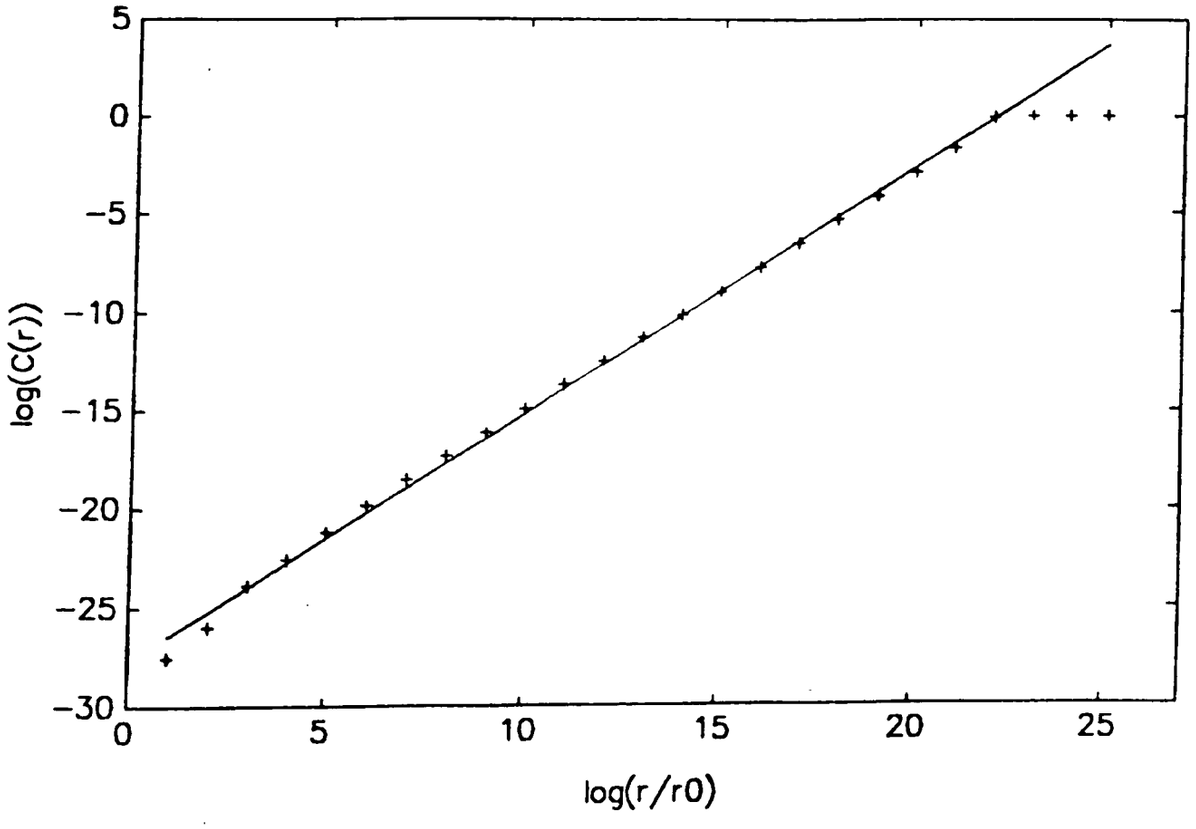


Figure 5.13 Log $C(r)$ versus $\log(r)$ for the chaotic attractor of the Hénon map.

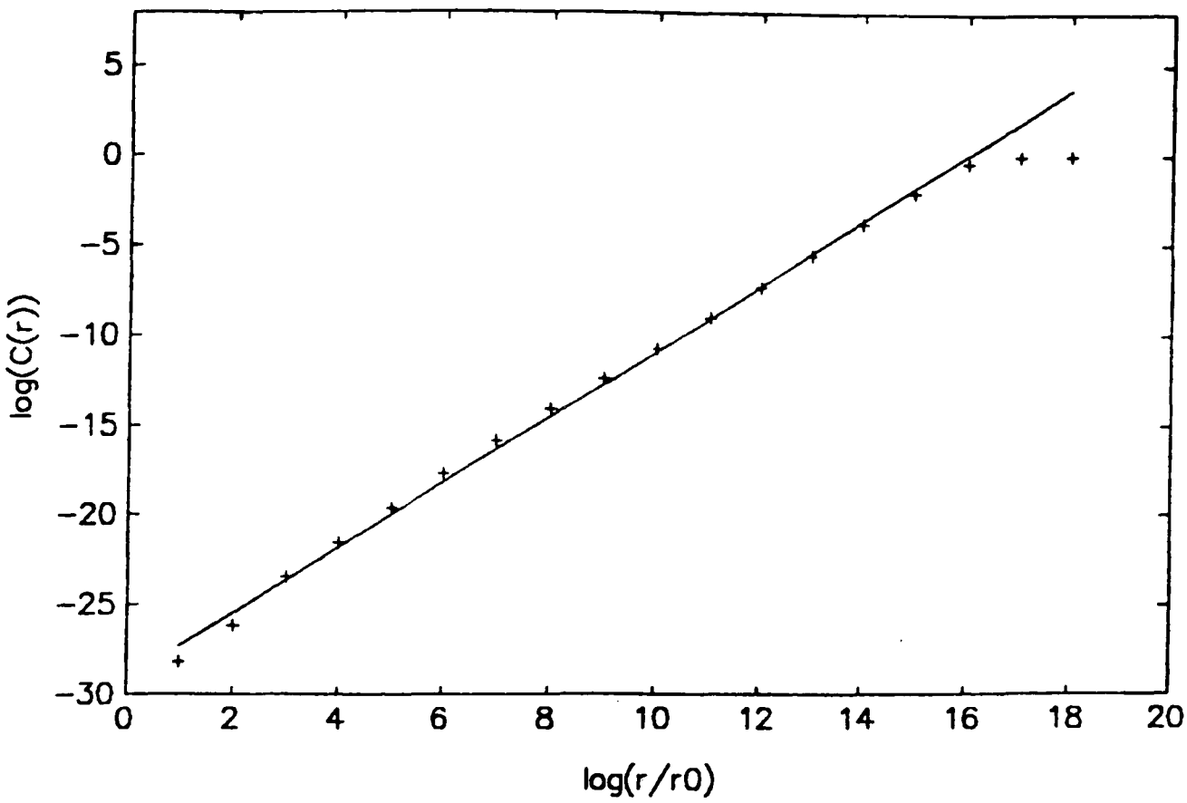


Figure 5.14 Log $C(r)$ versus $\log(r)$ for the chaotic attractor of the unconstrained or singly constrained gravity model.

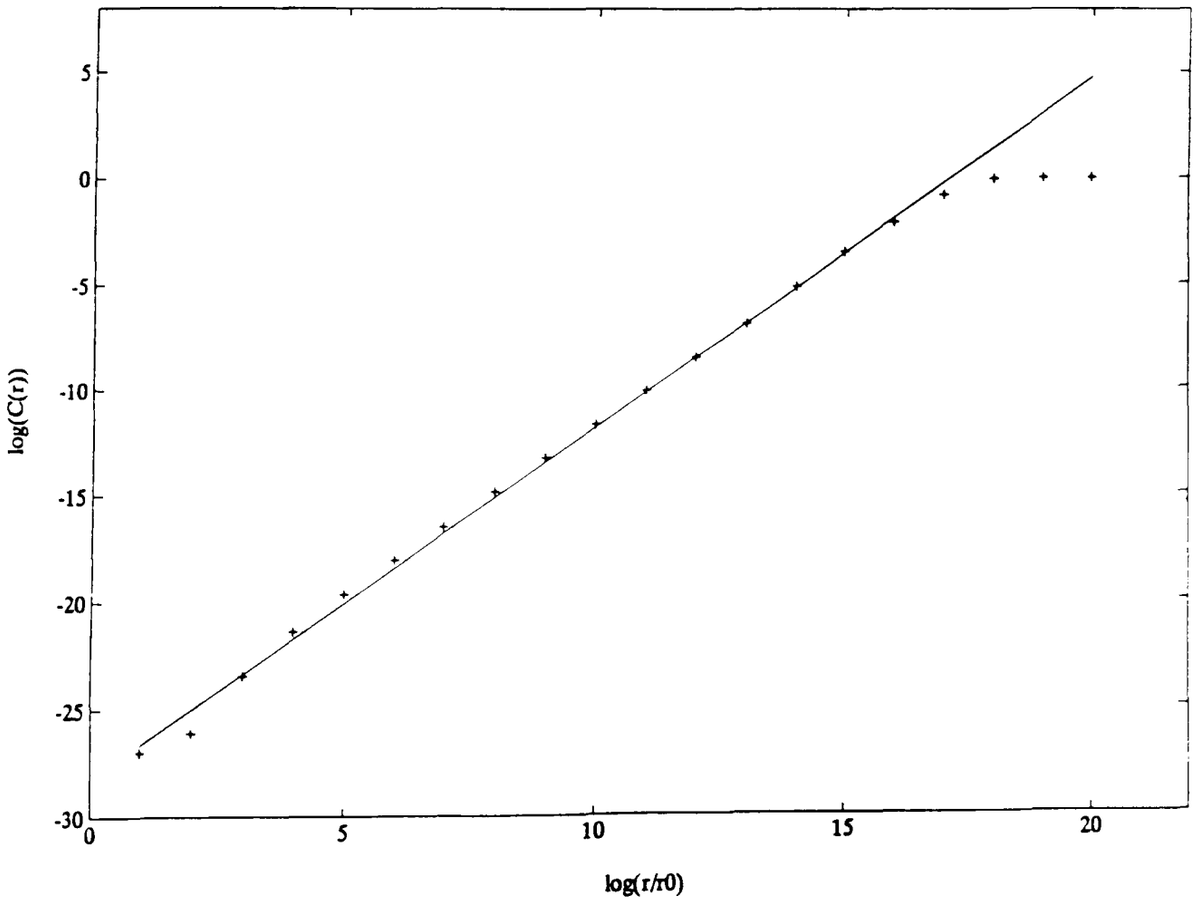


Figure 5.15 Log $C(r)$ versus $\log(r)$ for the chaotic attractor of the doubly constrained gravity model.

CHAPTER 6. THE DYNAMIC BEHAVIOUR OF THE LOGIT-BASED TRIP ASSIGNMENT MODEL

In this chapter, the variations of traffic flows in a road network are investigated based on the dynamic logit-based trip assignment model. Trip assignment has been one of the major subjects in traffic science. It is still an active research area. It concerns the formulation, analysis, and solution of large-scale models representing the complex interrelationships between flows and costs on the routes and links in a road network. A general road network can be very large and topologically complicated. Assigning O–D flows to such kind of network requires specially designed software packages, such as SATURN (Van Vliet, 1982) among others developed. The objective here, however, is to examine the dynamic logit-based assignment model so as to gain an understanding of the dynamic behaviour of the model.

6.1. INTRODUCTION

Given a trip matrix and a road network, trip assignment models allocate trips between each O–D pair to the routes connecting the O–D pair based on drivers' route choice behaviour. Before introducing the trip assignment model which will be examined, it is necessary to describe the notation that will be used.

6.1.1. Road network notation

A road network consists of a set of nodes, N, and a set of directed links L joining the nodes with each other. The set of nodes contains the origin nodes, the destination nodes, and intermediate nodes (intersections). The set of origin and destination nodes is denoted by M, where $M \subseteq N$. Each O–D pair (r,s) is connected by a set of routes (a chain of one or more links), p_{rs}. The set of all routes in the network is denoted by P, $P = \cup_{rs \in M} p_{rs}$. The topological

relationship of routes and links is represented by $\delta_{l,i}^{rs}$, where $\delta_{l,i}^{rs} = 1$ if link l is in the route i joining O–D pair (r,s), and $\delta_{l,i}^{rs} = 0$ otherwise.

The trip matrix is denoted by $\mathbf{t}$ with t_{rs} being the number of trips in a unit time period from origin r to destination s. Associated with each link there is a flow and a travel cost. Denote the flow on link l by y_l and the corresponding cost, which is a function of y_l by d_l. Then $d_l(y_l)$ is called the *link performance function* for link l. Similarly, for each route there is also a flow and a travel cost. The flow on each link is clearly the sum of the flows on those routes that use the link; while the cost on each route is assumed to be the sum of the costs on the links that form the route. Let the flow and the cost on route i joining O–D pair (r,s) be x_i^{rs} and c_i^{rs}, respectively. Then link flows and route flows are related by

$$y_l = \Sigma_{rs \in M} \Sigma_{i \in p_{rs}} \delta_{l,i}^{rs} x_i^{rs}, \quad l \in L.$$

The route costs and link costs are related by

$$c_i^{rs} = \Sigma_{l \in L} \delta_{l,i}^{rs} d_l, \quad i \in p_{rs}, \ rs \in M.$$

Sometimes vector notation is used for simplification. Given a particular order of the routes joining O–D pair (r,s), the vector of route flows on these routes is denoted by $\mathbf{x}^{rs}$. The vector of route flows on all routes in the road network is $(..., \mathbf{x}^{rs}, ...)$, which is denoted by $\mathbf{x}$. The notations for a road network are summarized in Table 6.1.

6.1.2. The model

The dynamic logit-based trip assignment model describes, for a given O–D matrix, the adjustments of the flow pattern in a road network from one time instant to another (for example, from day to day). The assignment at each time period is modelled by the logit model. It is assumed that a drivers' route choice is based on the travel costs in the previous time period and that the travel cost on each link depends on the flow on that link.

The flows assigned to each route must be nonnegative. They must also satisfy the O–D flow constraints. In other words, the sum of flows on all the routes joining an O–D pair must equal the number of trips between the O–D pair. Thus the flow x_i^{rs} on route i joining O–D pair (r,s) must satisfy

$$x_i^{rs} \geq 0, \qquad i \in p_{rs}, \ rs \in M, \tag{6.1a}$$

and

$$\Sigma_{i \in p_{rs}} x_i^{rs} = t_{rs}, \qquad rs \in M. \tag{6.1b}$$

Let U be the set of $\mathbf{x}$ that satisfies the above conditions. Then

$$U = \prod_{rs \in M} S^{rs}, \tag{6.2}$$

where

$$S^{rs} = \{(..., x_i^{rs}, ...): \ x_i^{rs} \geq 0, \ \Sigma_{i \in p_{rs}} x_i^{rs} = t_{rs}\}, \quad rs \in M,$$

so that

$$U = \{(..., \mathbf{x}^{rs}, ...): \ x_i^{rs} \geq 0, \ \Sigma_{i \in p_{rs}} x_i^{rs} = t_{rs}, \ rs \in M\}.$$

This set defines the phase space for the dynamic assignment model. The model can now be written as

$$\mathbf{G}: \ U \to U, \qquad G_i^{rs}(\mathbf{x}) = t_{rs} \frac{\exp(-\theta c_i^{rs})}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs})}, \quad i \in p_{rs}, \ rs \in M, \tag{6.3a}$$

where

$$c_i^{rs} = \Sigma_{l \in L} \delta_{l,i}^{rs} d_l(y_l), \quad i \in p_{rs}, \ rs \in M, \tag{6.3b}$$

$$y_l = \Sigma_{rs \in M} \Sigma_{i \in p_{rs}} \delta_{l,i}^{rs} x_i^{rs}, \quad l \in L, \tag{6.3c}$$

and θ is a positive parameter. The map $\mathbf{G}$ defines an adjustment mechanism in the dynamic assignment process: the flow pattern $\mathbf{x}$ at one time period will become $\mathbf{G}(\mathbf{x})$ at the next time period. Let n be the discrete time.

Then

$$\mathbf{x}(n+1) = \mathbf{G}(\mathbf{x}(n)), \qquad \mathbf{x}(n) \in U. \tag{6.4}$$

This assignment model has some similarities to the gravity model with the exponential deterrence function. In the assignment model, the current flow on route i joining O–D pair (r,s), $x_i^{rs}(n+1)$, depends on the previous costs on all the routes joining the O–D pair, c_j^{rs}, $j \in p_{rs}$. While each c_j^{rs} depends, in turn, on the costs on all the links in route j. It is very common for a network to have one particular link used by more than one route which may well join different O–D pairs. If there is no route overlapping another route connecting a different O–D pair (overlaps of routes between the same O–D pair are not ruled out), then (6.3b) and (6.3c) become

$$c_i^{rs} = \Sigma_{l \in L}\, \delta_{l,i}^{rs}\, d_l(\Sigma_{i \in p_{rs}}\, \delta_{l,i}^{rs}\, x_i^{rs}) \equiv c_i^{rs}(\mathbf{x}^{rs}).$$

Clearly in this case, $x_i^{rs}(n+1)$ will depend only on those $x_j^{rs}(n)$ with $j \in p_{rs}$. As a result, the assignment model (6.3) can be divided into several independent maps on S^{rs}, each of which may be written as

$$G_i^{rs}(\mathbf{x}) = t_{rs} \frac{\exp(-\theta c_i^{rs}(\mathbf{x}^{rs}))}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs}(\mathbf{x}^{rs}))} \equiv G_i^{rs}(\mathbf{x}^{rs}), \quad i \in p_{rs}.$$

Further, if there are no overlapping routes on any link at all, then the right hand side of (6.3c) will have only one term. In other words, the flows on all the links in each route are equal to the flow on that route. Consequently, c_i^{rs} will depend on x_i^{rs} only and the assignment model for each O–D pair may be written as

$$G_i^{rs}(\mathbf{x}^{rs}) = t_{rs} \frac{\exp(-\theta c_i^{rs}(x_i^{rs}))}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs}(x_j^{rs}))}, \quad i \in p_{rs}, \tag{6.5}$$

where

$$c_i^{rs}(x_i^{rs}) \equiv \Sigma_{l \in L} \delta_{l,i}^{rs} d_l(x_i^{rs}).$$

It can be seen that each component of the original model (6.3), that is, the assignment model for a single O–D pair, is now mathematically the same as the unconstrained gravity model with the exponential deterrence function. The only difference is that the cost c_i^{rs} here is the sum of several increasing functions of x_i^{rs} rather than a single function.

The assignment model will be examined in the similar way to that for the gravity model. The existence, uniqueness, and the stability of an equilibrium in the model are investigated theoretically in the next section. Numerical analysis is made in the subsequent section. The chapter is summarized in the last section.

6.2. THEORETICAL ANALYSIS

6.2.1. The existence of an equilibrium

The existence of a fixed point in the dynamic assignment model (6.3) can be assured by Brouwer's fixed point theorem mentioned in Chapter 5. The model defined by (6.3) is a continuous map on the closed convex set, U, so according to Brouwer's fixed point theorem it has at least one fixed point. A fixed point $\mathbf{x}^e \in U$ is given by $\mathbf{x}^e = \mathbf{G}(\mathbf{x}^e)$, or

$$x_i^{rs} = t_{rs} \frac{\exp(-\theta c_i^{rs})}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs})}, \quad i \in p_{rs}, \ rs \in M, \tag{6.6}$$

where

$$c_i^{rs} = \Sigma_{l \in L} \delta_{l,i}^{rs} d_l(y_l),$$
$$y_l = \Sigma_{rs \in M} \Sigma_{i \in p_{rs}} \delta_{l,i}^{rs} x_i^{rs}.$$

Here, the superscript "e" representing the equilibrium has been omitted. The equilibrium cannot be found analytically because it is the solution of a set of nonlinear equations. However, it can be obtained by numerical calculations.

6.2.2. The uniqueness of the equilibrium

Before considering the uniqueness of the equilibrium in the dynamic assignment model, it is important to review two notions of equilibrium discussed in the traffic assignment literature. They are the user equilibrium (UE), also called the Wardrop equilibrium, and the stochastic user equilibrium (SUE). In Chapter 2, it was pointed out that traffic assignment methods that consider congestion effects, normally by using link performance functions, are equilibrium assignment methods. These methods seek to find a flow pattern that satisfies the UE or SUE criteria. The majority of the researches consider steady-state (static) conditions and the equilibrium assignment problems are transformed to equivalent mathematical programming problems.

The *user equilibrium* is a flow pattern under which no driver can reduce his or her travel cost by changing routes. More specifically, at user equilibrium, for each O–D pair, the travel cost on all used routes is equal, and is not greater than the travel cost on any unused route if a driver would use that route.

Beckmann *et al.* (1956) formulated the problem of finding the user equilibrium as the following mathematical programming problem:

$$\text{Minimize} \qquad z(\mathbf{x}) = \Sigma_{l \in L} \int_0^{y_l} d_l(w)\, \mathrm{d}w \tag{6.7a}$$

subject to

$$\Sigma_{i \in p_{rs}} x_i^{rs} = t_{rs}, \quad rs \in M, \tag{6.7b}$$

$$x_i^{rs} \geq 0, \quad i \in p_{rs}, \quad rs \in M, \tag{6.7c}$$

$$y_l = \Sigma_{rs \in M} \Sigma_{i \in p_{rs}} \delta_{li}^{rs} x_i^{rs}, \quad l \in L. \tag{6.7d}$$

It is shown (Beckmann *et al.*, 1956) that a flow pattern that minimizes the objective function is identical to the user equilibrium and that there is only one minimum in the problem.

The UE assignment assumes that drivers have perfect knowledge of travel costs in the network and that all drivers react in the same way in their route choice. The stochastic equilibrium assignment, on the other hand, assumes that route

choice is based on the *perceived* travel cost rather than the real cost or the *measured* cost. The perceived cost is, in general, different for different drivers. *Stochastic user equilibrium* is a flow pattern under which no driver can reduce his or her perceived cost by changing a route. At the equilibrium, the real travel costs on all used routes between each O–D pair are not necessarily the same but they are such that each driver believes that his or her travel cost is minimized and cannot be improved by changing a route. As has been pointed out in Chapter 2 that the SUE assignment models are deterministic because the route or link flows used to define the SUE are deterministic variables.

One type of stochastic assignment method is the logit-based model. It has been shown by Fisk (1980) that the logit-based stochastic user equilibrium is the same as the unique solution of the following convex programming problem

$$\text{Minimize} \qquad z(\mathbf{x}) = \frac{1}{\theta}\Sigma_{rs\in M}\Sigma_{i\in p_{rs}} x_i^{rs} \ln x_i^{rs} + \Sigma_{l\in L}\int_0^{y_l} d_l(w)\,\mathrm{d}w \tag{6.8a}$$

subject to

$$\Sigma_{i\in p_{rs}} x_i^{rs} = t_{rs}, \quad rs\in M, \tag{6.8b}$$

$$x_i^{rs} \geq 0, \quad \forall i\in p_{rs}, \quad rs\in M, \tag{6.8c}$$

$$y_l = \Sigma_{rs\in M}\Sigma_{i\in p_{rs}} \delta_{l,i}^{rs} x_i^{rs}, \quad l\in L. \tag{6.8d}$$

The solution to the problem can be written implicitly as (Fisk, 1980)

$$x_i^{rs} = t_{rs}\frac{\exp(-\theta c_i^{rs})}{\Sigma_{j\in p_{rs}}\exp(-\theta c_j^{rs})}, \quad i\in p_{rs}, \quad rs\in M, \tag{6.9}$$

where

$$c_i^{rs} = \Sigma_{l\in L}\,\delta_{l,i}^{rs}\,d_l(y_l),$$
$$y_l = \Sigma_{rs\in M}\Sigma_{i\in p_{rs}}\,\delta_{l,i}^{rs}\,x_i^{rs}.$$

Further, if link costs are constant (flow independent), the flow pattern defined by the solution will be identical to that produced by the static logit-based

assignment model first suggested by Dial (1971). Dial's model may be written as

$$x_i^{rs} = t_{rs} \frac{\exp(-\theta c_i^{rs})}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs})}, \quad i \in p_{rs}, \ rs \in M,$$

where

$$c_i^{rs} = \Sigma_{l \in L} \delta_{l,i}^{rs} d_l$$

Note that here d_l is independent of y_l. As $\theta \to \infty$, the importance of the second term of the objective function becomes dominant and the solution tends to the user equilibrium (the solution to (6.7)). See Fisk (1980) for detailed discussions.

It can be seen that the fixed point (6.6) in the dynamic model (6.3) is exactly the same as the solution (6.9) to Fisk's program (6.8). The phase space (6.2) of the dynamic assignment model is convex and is the same as the feasible region defined by (6.8b–6.8c) in this program. Therefore, the equilibrium in the dynamic assignment model is unique as well. The optimum solution (6.9) gives the stochastic user equilibrium (SUE) assignment, so does the fixed point in the dynamic assignment model. Both the flow pattern by Dial's model and user equilibrium are included as special cases of the equilibrium of the dynamic model.

Being unique and giving the SUE, the equilibrium in the dynamic model is rather favourable. An immediate concern is its stability, which is considered next.

6.2.3. The stability of the equilibrium

There have been relatively few studies on the stability of an equilibrium in trip assignment, partly because of the lack of suitable dynamic models. Equilibrium assignment models that consider only a steady state assume implicitly that the equilibrium is stable. To consider the stability of an equilibrium, it is necessary to model the adjustment of flow patterns over time. The studies by Horowitz (1984) and Smith (1984) described in Chapter 2 are in this line. In Horowitz (1984) the stability of an equilibrium in a discrete model for a network of one O–D pair connected by two links is considered, while in Smith (1984) the stability of user equilibrium in a continuous model is analyzed. Here, the stability of the equilibrium in the dynamic model (6.3) will be investigated.

Consider, first, a simple case in which there is one O–D pair joined by two routes with no common link. In (6.5), putting $t_{rs} = 1$ and using subscripts only, then the model becomes

$$G_i(\mathbf{x}) = \frac{\exp(-\theta c_i(x_i))}{\Sigma_j \exp(-\theta c_j(x_i))}, \quad i = 1, 2, \tag{6.10}$$

where

$$c_i(x_i) = \Sigma_{l \in L} \delta_{l,i} d_l(x_i), \quad i = 1, 2,$$

$$\delta_{l,i} = \begin{cases} 1 & \text{if link } l \text{ is in route } i \\ 0 & \text{otherwise} \end{cases}$$

and $d_l(y_l)$ is the link performance function for link l. The equilibrium in this model can be written as

$$x_i^e = \frac{\exp(-\theta c_i(x_i^e))}{\Sigma_j \exp(-\theta c_j(x_i^e))}, \quad i = 1, 2,$$

where

$$c_i(x_i^e) = \Sigma_{l \in L} \delta_{l,i} d_l(x_i^e), \quad i = 1, 2.$$

Since model (6.10) is mathematically the same as the one-dimensional gravity model with the exponential deterrence function, the results of the stability analysis of the gravity model in section 5.2.4 can be applied here. In (5.9), setting $\mu = 0$, replacing β by θ, t_i^e by x_i^e, and $c_i'(t_i^e)$ by $c_i'(x_i^e)$, the condition for the equilibrium in model (6.10) to be stable is

$$\theta x_1^e x_2^e [c_1'(x_1^e) + c_2'(x_2^e)] < 1, \tag{6.11a}$$

where

$$c_i'(x_i^e) = \Sigma_{l \in L} \delta_{l,i} \frac{\mathrm{d}}{\mathrm{d}x_i} [d_l(x_i^e)], \quad i = 1, 2.$$

It needs to be borne in mind here that $c_i'(x_i^e) \geq 0$ because the link performance functions are normally assumed to be increasing functions. Similarly, from (5.10b), a sufficient condition for the equilibrium to be stable can be given by

$$\theta\,[c_1'(x_1^e) + c_2'(x_2^e)] < 4, \tag{6.11b}$$

where

$$c_i'(x_i^e) = \Sigma_{l\in L}\,\delta_{l,i}\frac{\mathrm{d}}{\mathrm{d}x_i}[d_l(x_i^e)], \quad i = 1, 2.$$

Also from the analysis of the one-dimensional gravity model in 5.2.4 it can be concluded that the fixed point and period–2 orbits are the only possible attractors that could occur in model (6.10); any trajectory in the phase space will be attracted to the stable fixed point, or a stable period–2 orbit if the fixed point is unstable.

The stability of the equilibrium in the general model (6.3) is investigated in the same way as that used in the last chapter. The unique equilibrium in (6.3) is locally asymptotically stable if the magnitude of all eigenvalues of the Jacobian matrix of $\mathbf{G}$ at the equilibrium is less than 1 (Parker and Chua, 1989). The eigenvalues cannot be evaluated analytically. A sufficient condition for the stability of the equilibrium can be obtained by bounding the eigenvalues of the matrix by the p–norm, where $p = 1$, based on the theorem used in section 5.2.4.

Denote the Jacobian matrix of $\mathbf{G}$ evaluated at the equilibrium $\mathbf{x}^e$ by $\mathbf{D}$. Then from (6.3), the elements of the matrix are given by

$$D_{ij} = \frac{\partial}{\partial x_j^{uv}} G_i^{rs}(\mathbf{x})$$

$$= \theta\, x_i^{rs}\left[-\frac{\partial c_i^{rs}}{\partial x_j^{uv}} + \frac{1}{t_{rs}}\Sigma_{k\in p_{rs}}\, x_k^{rs}\frac{\partial c_k^{rs}}{\partial x_j^{uv}}\right], \quad i\in p_{rs},\ j\in p_{uv},\ rs,uv\in M,$$

where

$$\frac{\partial c_i^{rs}}{\partial x_j^{uv}} = \sum_{l\in L}\delta_{l,i}^{rs}\,(d_l'(y_l))\left[\frac{\partial y_l}{\partial x_j^{uv}}\right]$$

$$= \sum_{l\in L}\delta_{l,i}^{rs}\,\delta_{l,j}^{uv}\,(d_l'(y_l)).$$

Here, the superscripts "e" implying the equilibrium are omitted and will be throughout the rest of the section. The superscripts uv are used to signify that the routes i and j do not necessarily join the same O–D pair. The 1–norm for

the matrix is

$$\|\mathbf{D}\|_1 = \max_{j\in P} \Sigma_{rs\in M} \Sigma_{i\in p_{rs}} |D_{ij}|$$

$$= \max_{j\in P} \Sigma_{rs\in M} \Sigma_{i\in p_{rs}} \theta\, x_i^{rs} \left| -\frac{\partial c_i^{rs}}{\partial x_j^{uv}} + \frac{1}{t_{rs}} \Sigma_{k\in p_{rs}} x_k^{rs} \frac{\partial c_k^{rs}}{\partial x_j^{uv}} \right|.$$

According to the theorem in section 5.2.4, none of the eigenvalues of the Jacobian matrix has a magnitude greater than the above norm. A sufficient condition for the stability of the equilibrium can therefore be given by

$$\max_{j\in P} \Sigma_{rs\in M} \Sigma_{i\in p_{rs}} \theta\, x_i^{rs} \left| -\frac{\partial c_i^{rs}}{\partial x_j^{uv}} + \frac{1}{t_{rs}} \Sigma_{k\in rs} x_k^{rs} \frac{\partial c_k^{rs}}{\partial x_j^{uv}} \right| < 1. \tag{6.12}$$

In the special case where there is not any route overlapping, we have

$$\frac{\partial c_i^{rs}}{\partial x_j^{uv}} = \Delta_{ij} \frac{\partial c_j^{uv}}{\partial x_j^{uv}} = \Delta_{ij} \Sigma_{l\in L}\, \delta_{l,j}^{uv} \frac{\mathrm{d}d_l}{\mathrm{d}y_l},$$

where

$$\Delta_{ij} = \begin{cases} 0 & i \neq j \\ 1 & i = j, \end{cases}.$$

Note that $i = j$ implies $rs = uv$. So

$$\frac{\partial}{\partial x_j^{uv}} G_i^{rs}(\mathbf{x}) = \theta\, x_j^{uv} \frac{\partial c_j^{uv}}{\partial x_j^{uv}} \left[-\Delta_{ij} + \frac{x_i^{uv}}{t_{uv}} \right]$$

if routes i and j both join the same O–D pair, and

$$\frac{\partial}{\partial x_j^{uv}} G_i^{rs}(\mathbf{x}) = 0$$

otherwise. The norm now becomes

$$\|\mathbf{D}\|_1 = \max_{j\in P} \Sigma_{i\in p_{rs}} \theta\, x_j^{uv} \frac{\partial c_j^{uv}}{\partial x_j^{uv}} \left| - \Delta_{ij} + \frac{x_i^{uv}}{t_{uv}} \right|$$

$$= \max_{j\in P} 2\ \theta\, x_j^{uv} \frac{\partial c_j^{uv}}{\partial x_j^{uv}} \left[1 - \frac{x_j^{uv}}{t_{uv}} \right]$$

$$= \max_{j\in P} 2\ \theta\, t_{uv} \frac{\partial c_j^{uv}}{\partial x_j^{uv}} \frac{x_j^{uv}}{t_{uv}} \left[1 - \frac{x_j^{uv}}{t_{uv}} \right].$$

Since $\frac{x_j^{uv}}{t_{uv}} + \left[1 - \frac{x_j^{uv}}{t_{uv}} \right] = 1$, we have $\frac{x_j^{uv}}{t_{uv}} \left[1 - \frac{x_j^{uv}}{t_{uv}} \right] \leq 0.25.$

Therefore,

$$\|\mathbf{D}\|_1 \leq \max_{j\in P} 0.5\ \theta\, t_{uv} \frac{\partial c_j^{uv}}{\partial x_j^{uv}}.$$

The sufficient condition for the stability of the equilibrium is then reduced to

$$\max_{j\in P} \theta\, t_{uv} \frac{\partial c_j^{uv}}{\partial x_j^{uv}} < 1/0.5 = 2. \tag{6.13}$$

Clearly this condition is similar to the sufficient condition (5.11b) for the stability of the equilibrium in the gravity model with the exponential deterrence function. See section 5.2.4. This similarity agrees with the fact that when there is no route overlapping, the trip assignment model (6.3) is mathematically the same as the gravity model with the exponential deterrence function.

From the conditions (6.11) to (6.13) it can be seen that the stability of the equilibrium in the dynamic model depends on the values of parameter θ and the derivatives of the link performance functions at the equilibrium. Clearly, the larger θ is, the less the chance is for the equilibrium to be stable. In conditions (6.11) and (6.13), where there are no overlapping routes in the network, the larger the derivative of the link performance function is, the less likely that the

equilibrium is stable. This is because larger derivatives imply stronger dependence of link costs on link flows so that small variations of link flows can cause large changes in link costs and so in link flows at the next time period. The situation in which there are overlapping routes is more complicated and is investigated numerically in the next section.

6.3. NUMERICAL ANALYSIS

In this section, the assignment model is studied numerically for other possible attractors in addition to the equilibrium. Given a road network and a trip matrix, numerical analysis of the model involves iterating the equation (6.4)

$$\mathbf{x}(n+1) = \mathbf{G}(\mathbf{x}(n)), \qquad \mathbf{x}(n) \in U.$$

until a steady state is reached. At each iteration, the link flows are calculated first from the route flows by

$$y_l = \Sigma_{rs \in M} \Sigma_{i \in p_{rs}} \delta^{rs}_{l,i} x^{rs}_i, \quad l \in L,$$

from which the link costs are obtained by the link performance functions. Then the route costs are calculated by

$$c^{rs}_i = \Sigma_{l \in L} \delta^{rs}_{l,i} d_l(y_l), \quad i \in p_{rs}, \ rs \in M.$$

Finally a new set of route flows is obtained from

$$G^{rs}_i(\mathbf{x}) = t_{rs} \frac{\exp(-\theta c^{rs}_i)}{\Sigma_{j \in p_{rs}} \exp(-\theta c^{rs}_j)}, \quad i \in p_{rs}, \ rs \in M.$$

The iterations may continue as long as desired.

Two commonly used link performance functions (Branston, 1976) are considered. The first one is the BPR (Bureau of Public Roads) function, or the power function. The BPR function is one of the most widely used link performance functions. It may be written as

$$d_l(y_l) = d_{l0}[1+\alpha(y_l/q_l)^{\gamma}], \tag{6.14}$$

where d_{l0} is uncongested travel cost, q_l is link capacity, α and γ are positive constants. The second link performance function is the exponential function

$$d_l(y_l) = d_{l0}\ \alpha^{\gamma(y_l/q_l)}, \tag{6.15}$$

where d_{l0} is uncongested travel cost, q_l is link capacity, α and γ are positive constants, $\alpha > 1$.

It has been pointed out at the beginning of the chapter that assigning O–D flows to a general road network involves large amount of calculations and requires special software packages such as SATURN (Van Vliet, 1982). Therefore, here, some simple artificial networks are used for numerical calculations for the purpose of finding out typical dynamic behaviour in the assignment model.

Several networks are tested, using the two link performance functions and various values of parameters and initial conditions. In all the calculations made, only point attractors and period–2 attractors are found in the model. This outcome is not too surprising because it is consistent with the numerical analysis of the gravity model with the exponential deterrence function in Chapter 5. As has been mentioned, the two models are similar.

In the rest of this section, the results of the calculations for two networks with the two link performance functions are demonstrated; some typical values of parameters used in practice are examined to see if the equilibrium is stable when these values are used. Instead of showing the bifurcation diagrams which will all look very much the same as Figure 5.1 (the bifurcation diagram for the gravity model with the power deterrence function), the behaviour in the assignment model is summarized in the diagrams of attractor regimes in the parameter space. There are three parameters in the assignment model, θ, α, and γ. The diagrams will therefore be three dimensional, showing regimes of point attractors and period–2 attractors.

Computations for Network 1

Figure 6.1 shows the first network considered (Potts and Oliver, 1972, page 72). This network has one O–D pair, nine routes, and ten links. The chain of links for each route is listed in Table 6.2. The capacities and the uncongested travel costs on all links are shown in Table 6.3. The O–D flow $t_{11} = 9$.

Figures 6.2 and 6.3 show the attractor regimes in the parameter space calculated for this network with the two link performance functions. In these diagrams the region beneath the surface is the point attractor regime and the region above the surface is the period–2 attractor regime. The curves in the (α, γ)-plane are contour lines of the surface. It can be seen that for smaller values of parameters, the equilibrium in the model is stable. When the values of parameters become larger, the point attractors change into period–2 orbits.

Computations for Network 2

Network 2 is shown in Figure 6.4. This network is from Potts and Oliver (1972, page 113) but is modified slightly by adding one link (link 7) and by removing one node. The modified network has 4 O–D pairs, 12 routes, and 11 links. The topological structure of the network is summarized in Table 6.4 and the link data in Table 6.5. The trip matrix for the network is given in Table 6.6.

The attractor regimes for network 2 with the two link performance functions are shown in Figures 6.5 and 6.6 respectively, where the region underneath the surface is the point attractor regime and the region above the surface is the period–2 attractor regime. It can be seen that the diagrams for this network are very much the same as that for network 1, although the two networks are completely different. This suggests that the behaviour described so far is very typical in this trip assignment model.

Some practical considerations

Naturally, it is desirable to know if the equilibrium would be stable when commonly used values of parameters are used. In the BPR link performance function, typical parameter values are $\alpha = 0.15$ and $\gamma = 4.0$ (Branston, 1976),

respectively. While in the exponential link performance function, $\gamma = 1.0$ and α ranges from 1.0 to 2.0 (Branston, 1976).

The parameter θ in the logit model is related to the standard deviation of the distribution of perceived route costs. Larger values of θ correspond to smaller variations in the perceived route costs and so in the route choice. In the limit, when $\theta \to \infty$, all drivers' perceived costs are the same, and they choose the cheapest route (in terms of measured cost). On the other hand, when $\theta \to 0$, every route is chosen with the same probability and the flow between each O–D pair is equally shared among the routes between the O–D pair, regardless of measured route costs. It is therefore important to consider different values of θ. In the following calculations, different values of θ are considered. Calculations are made for the above two networks, using the typical values of parameters in the two link performance functions.

When the BPR function is used, the calculations showed that the equilibrium in network 1 is stable if $\theta < 0.45$ and that the equilibrium in network 2 is stable if $\theta < 0.15$. The critical value of θ may depend on how congested the network is. In network 1, if the O–D flow is 12 rather than 9, calculations showed that the equilibrium loses its stability at $\theta = 0.15$.

With the exponential performance function, two parameters are variable, α and θ. The attractor regimes in the $\alpha - \theta$ plane for the two networks are shown in Figures 6.7 and 6.8, respectively. The region below the line is the regime in which the equilibrium is stable. In these two diagrams, the region for the equilibrium in network 1 to be stable is larger than that in network 2. This is consistent with the result of the BPR link performance function.

In the dynamic assignment model, the relationships between the flows and the costs in the network affect the dynamic behaviour of the model in a similar way to that in the gravity model. Consider first the situation where there are no overlapping routes. If, at one time period, the flows on some routes or links are smaller, these routes or links will become cheaper to travel and will attract more flows at the next time period. If the values of parameters are small so that the compensation is mild, the dynamic process will converge to the equilibrium. If, on the other hand, the values of parameters are too large and the differences in route costs are over-compensated, oscillations will occur and may not die down.

The mechanism is similar in the case where there are overlapping routes because the logit assignment model treats overlapping routes as independent routes. If two or more routes overlap on a link, then this link tends to be heavily loaded. Suppose this is the case at some stage, then the overlapping routes will be very expensive and a large part of the flows will be diverted to alternative routes at the next stage. It is clear that the existence of overlapping routes tends to make oscillations of flows more serious. It has long been known that the logit-based model may produce unreasonable flows on overlapping routes (Sheffi, 1985). One way to overcome the deficiency is to consider the dependence of link costs on link flows. This is just what the dynamic model accommodates. However, the dynamic process can approach the equilibrium only if the values of parameters are not too large.

6.4. SUMMARY AND COMMENTS

The dynamic logit-based trip assignment model has been investigated theoretically and numerically. The equilibrium in the model is found to be unique and is the same as the optimal solution of the Fisk's mathematical programming problem, which gives the logit-based SUE in a road network. Both the flow pattern given by Dial's model and Wardrop's user equilibrium are special cases of the equilibrium in the dynamic model. A sufficient condition for the equilibrium to be stable is established. If the values of parameters are small so that the flows and costs vary moderately, the dynamic process will converge to the equilibrium. If, however, the values of parameters are too large so that the flows and costs depend on each other too much, the flows may oscillate perpetually and never approach the equilibrium. Empirical studies are needed to determine, for a particular road network, the values of parameters in link performance functions and the route choice functions (logit model) so that one can know if the equilibrium will be stable. Empirical studies may also be used to validate the model to see if the model is realistic or not.

The assignment model considered in this chapter has some advantages. It accounts for both variations in drivers' route choice behaviour and congestion effect. It has a unique equilibrium which is identical to the logit-based stochastic user equilibrium. Meanwhile, it is dynamic and so allows for stability analysis. It seems plausible to assume that drivers choose their routes, say, today based on the travel costs they experienced yesterday. Apparently, this may not be the

only adjustment mechanism of the flow pattern over time. Dynamic traffic assignment is still an active subject, particularly the type of dynamic considerations used here. The general problem of dynamics and stability of trip assignment is difficult. One reason is the lack of knowledge of how current route choices depend on past costs and how costs vary with flows, particularly when congestion effects are involved. The other reason is that even very simple relationships between network flows and costs may produce highly nonlinear dynamic models.

Table 6.1 Road network notation

N	set of nodes
L	set of links
M	set of origin and destination nodes; $M \subseteq N$
P	set of all routes
p_{rs}	set of routes connecting origin r and destination s; $p_{rs} \subseteq P$; $r, s \in M$
t_{rs}	total flow from origin r to destination s; $\mathbf{t} = (..., t^{rs}, ...)$
x_i^{rs}	flow on route i joining O–D pair (r,s)
$\mathbf{x}^{rs}$	vector of route flows on all routes joining O–D pair (r,s); $\mathbf{x}^{rs} = (..., x_i^{rs}, ...)$
$\mathbf{x}$	vector of route flows on all routes in the network; $\mathbf{x} = (..., \mathbf{x}^{rs}, ...)$
c_i^{rs}	cost on route i connecting O–D pair (r,s)
y_l	flow on link l
d_l	cost on link l
$d_l(y_l)$	link performance function for link l
$\delta_{l,i}^{rs}$	link-route indicator: $\delta_{l,i}^{rs} = \begin{cases} 1 \text{ if link } l \text{ is in route } i \text{ between O–D pair } (r,s) \\ 0 \text{ otherwise} \end{cases}$

Table 6.2 The structure of network 1

O–D pair	Route number	link chain
O_1–D_1	1	1,3,6,8,10
	2	1,3,6,9
	3	1,3,7,10
	4	1,4,8,10
	5	1,4,9
	6	1,5,10
	7	2,6,8,10
	8	2,6,9
	9	2,7,10

Table 6.3 The link data of network 1

link number	link capacity	link cost
1	6	1
2	4	5
3	4	1
4	6	3
5	1	7
6	3	1
7	4	6
8	6	3
9	4	6
10	6	3

Table 6.4 The structure of network 2

O–D pair & flow	Route number	link chain
O_1–D_1 (35)	1	1,6,8,10
	2	2,4,6,8,10
	3	2,5,9,10
O_1–D_2 (20)	1	1,6,8,11
	2	2,4,6,8,11
	3	2,5,9,11
O_2–D_1 (15)	1	3,4,6,8,10
	2	3,5,9,10
	3	7,9,10
O_2–D_2 (30)	1	3,4,6,8,11
	2	3,5,9,11
	3	7,9,11

Table 6.5 The link data of network 2

link number	link capacity	link cost
1	20	3
2	30	1
3	20	5
4	20	1
5	25	5
6	40	2
7	20	7
8	40	1
9	45	1
10	40	1
11	43	3

Table 6.6 Trip matrix for network 2

		destination		total
origin		1	2	$\Sigma_{j} t_{ij}$
1		35	20	55
2		15	30	45
total	$\Sigma_{i} t_{ij}$	50	50	$\Sigma_{ij} t_{ij}$=100

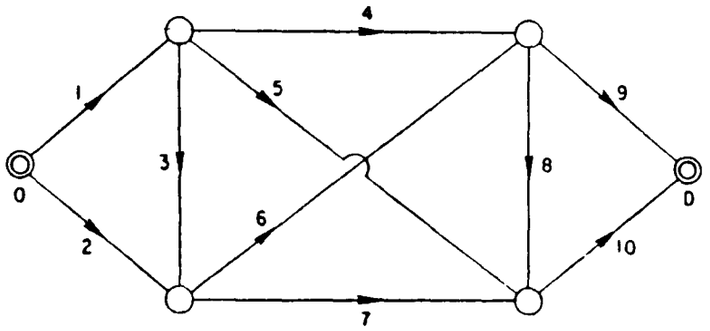


Figure 6.1 Road network 1 (Potts and Oliver, 1972, page 72).

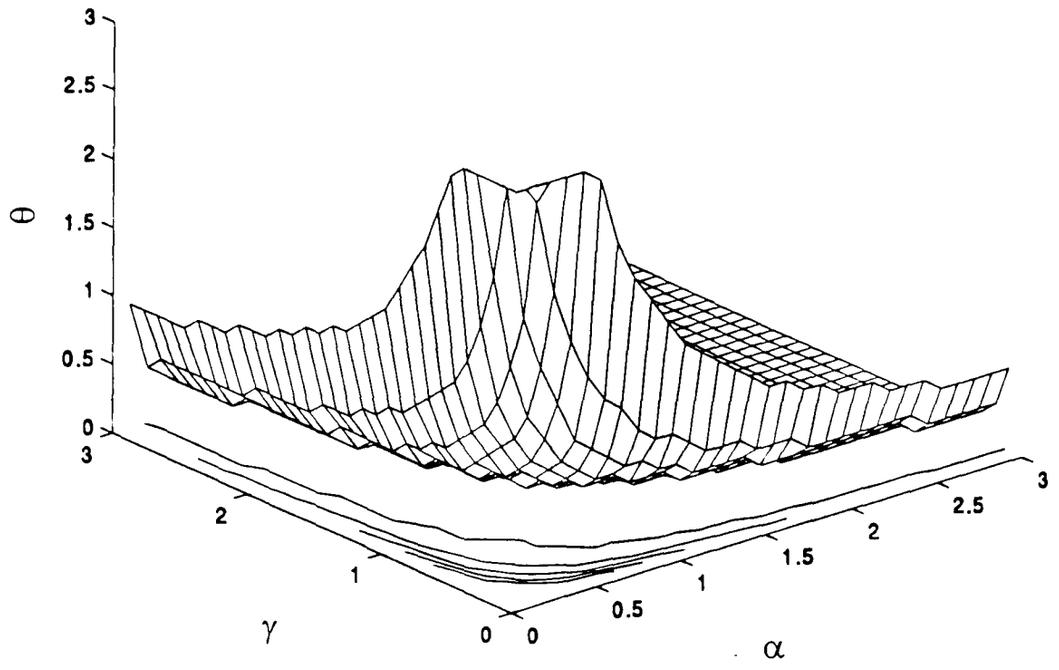


Figure 6.2 Attractor regimes in the parameter space for the assignment model for network 1, with the BPR link performance function.

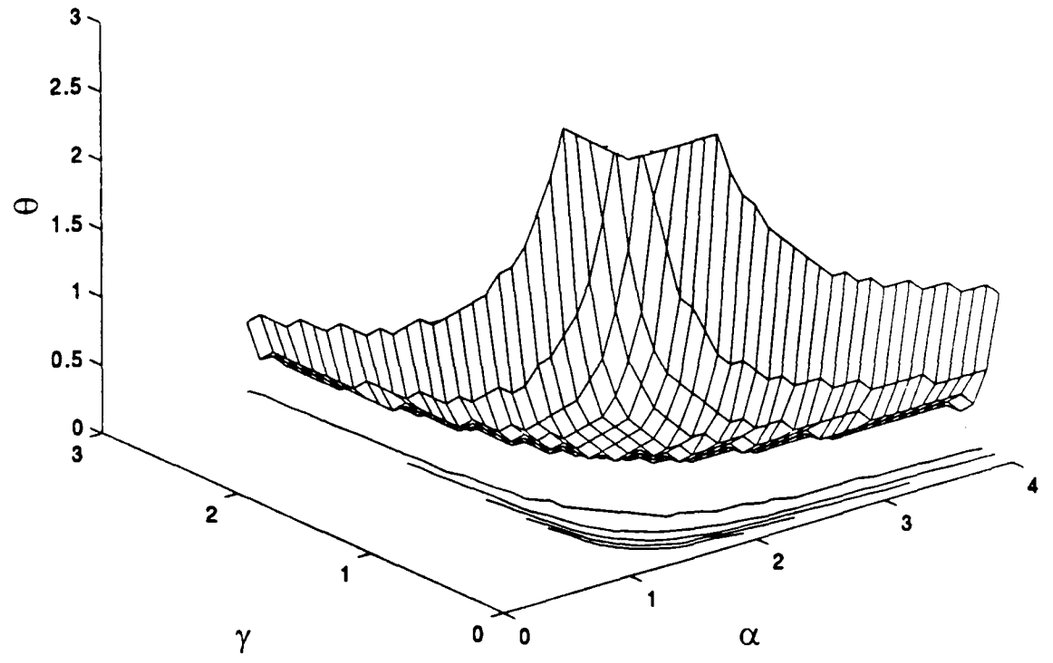


Figure 6.3 Attractor regimes in the parameter space for the assignment model for network 1, with the exponential link performance function.

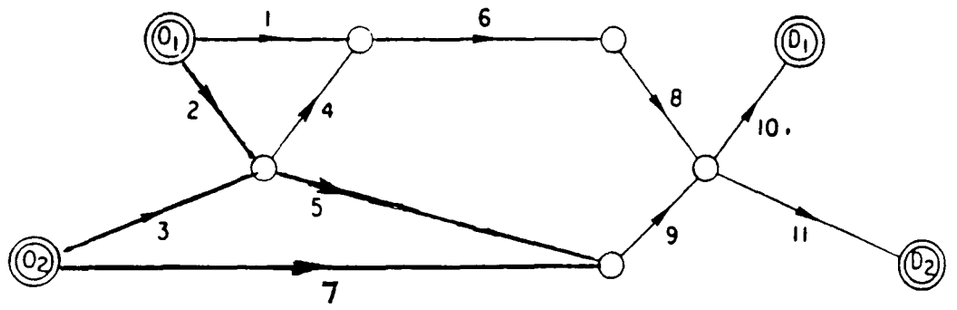


Figure 6.4 Road network 2 (modified from Potts and Oliver, 1972, page 113).

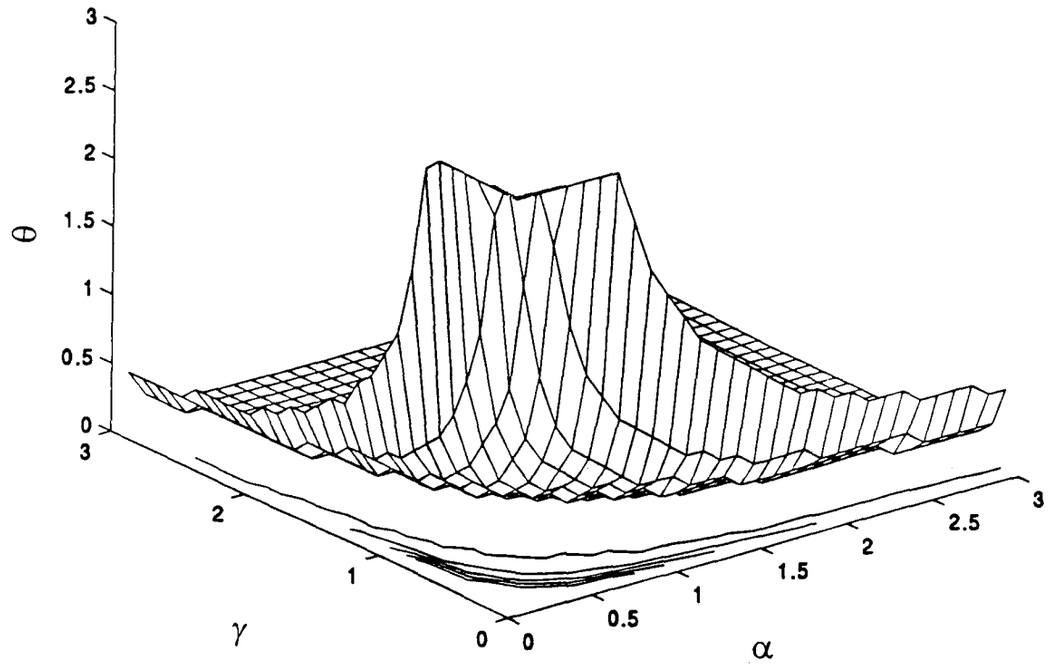


Figure 6.5 Attractor regimes in the parameter space for the assignment model for network 2, with the BPR link performance function.

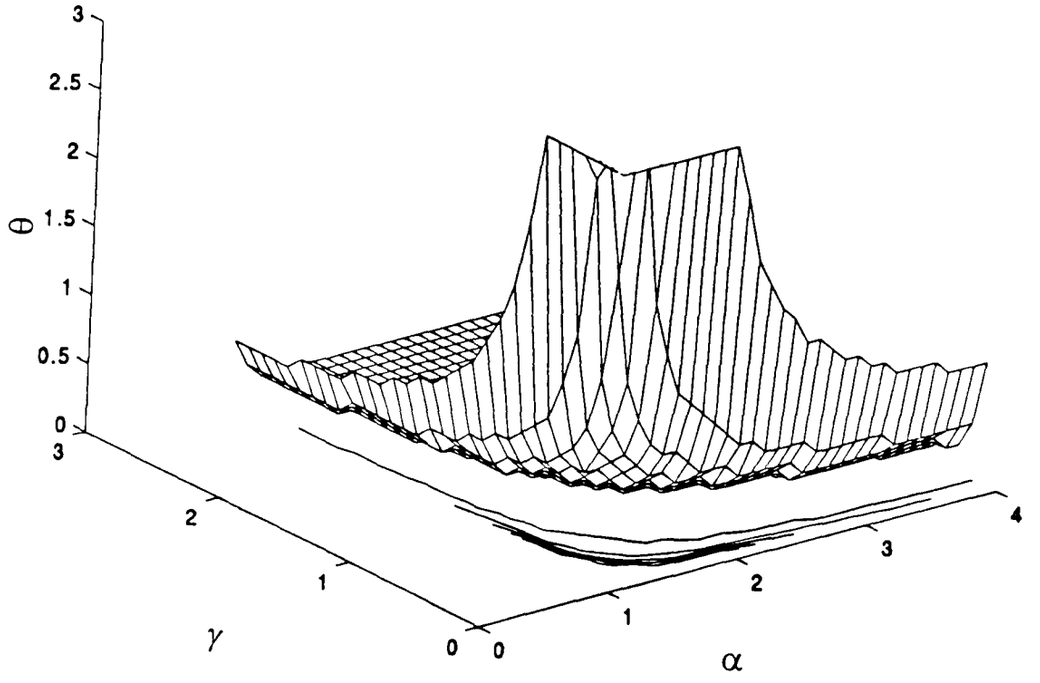


Figure 6.6 Attractor regimes in the parameter space for the assignment model for network 2, with the exponential link performance function.

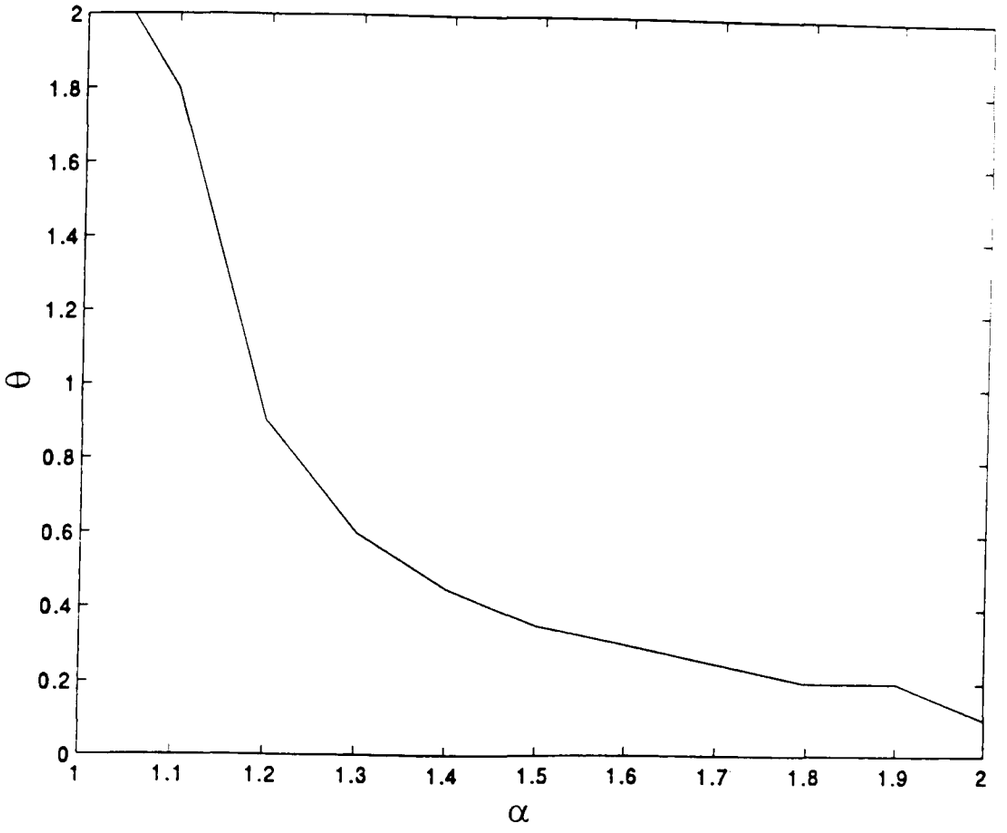


Figure 6.7 Attractor regimes in the $\alpha - \theta$ plane for the assignment model for network 1, with the exponential link performance function.

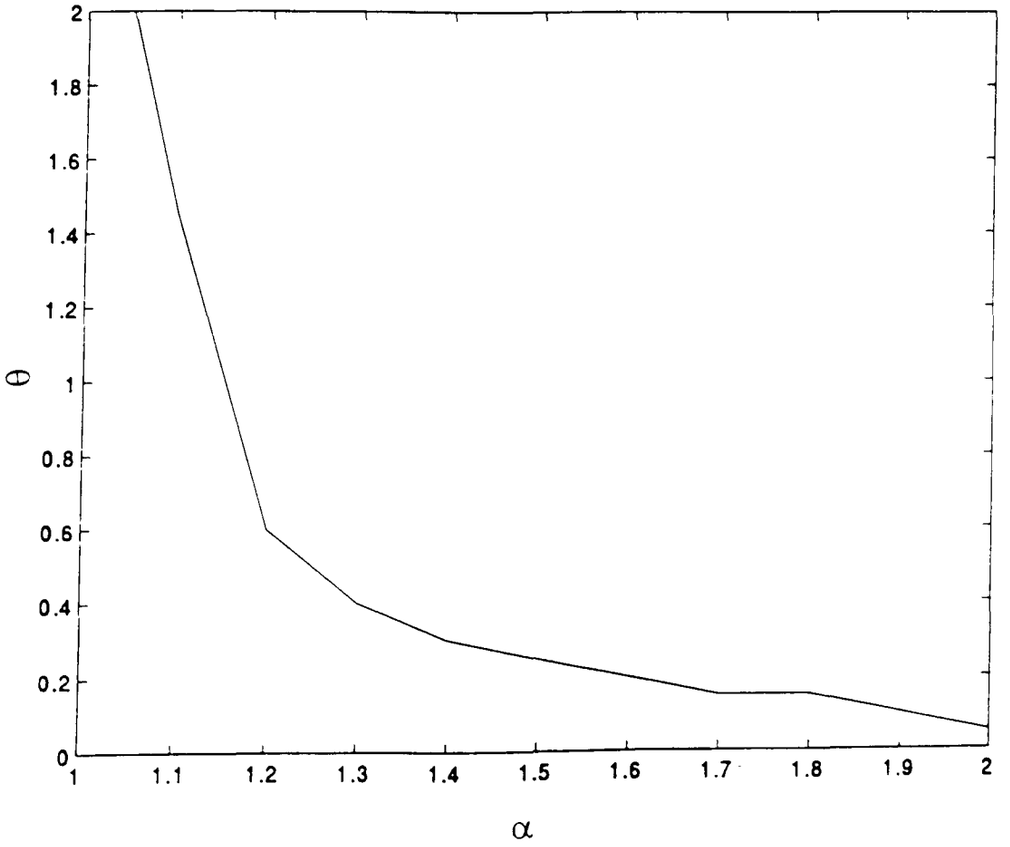


Figure 6.8 Attractor regimes in the $\alpha - \theta$ plane for the assignment model for network 2, with the exponential link performance function.

CHAPTER 7. CONCLUSIONS

The aim of this thesis was to investigate the time variation of traffic characteristics, including the speeds of cars and the spacing between the cars on a road link, the flows on the routes or links in a road network, and the flows between O–D pairs in an O–D network. Traffic flows on road links and in road networks have been treated as dynamical systems modelled by differential or difference equations. In this final chapter, the dynamic behaviour found in the three traffic models is summarized; some limitations and possible extensions of the present research are discussed.

7.1. TRAFFIC DYNAMICS REVEALED

The dynamic behaviour of road traffic flows both at the link level and at the network level has been investigated based on the car-following model, the gravity model, and the logit-based trip assignment model. The behaviour of the three models is summarized in turn first. Some general conclusions and practical implications follow next.

7.1.1. The car-following model

The car-following model is a system of delay-differential equations describing a line of cars moving on a single lane of a road link. The model assumes that each driver responds to the variation in the motion of the car in front. The model was investigated under two assumptions about the motion of the first car. The first assumption is that the first car moves at a constant speed; the model is then autonomous. The second assumption is that the speed of the first car is a sinusoidal wave; the motion of the first car then serves as an input disturbance. The car-following model is non-autonomous in this case.

The stability of the linear car-following model was analysed theoretically: the

local stability of the equilibrium in the nonlinear autonomous model was investigated by linearization. The car-following model was integrated numerically for various values of parameters in an attempt to identify all possible types of dynamic behaviour in the nonlinear model.

Typical behaviour in the autonomous model was found to be an equilibrium where the speed of each car is the same and is equal to the speed of the first car, and the distances between the cars are constant, not necessarily equal to each other. The equilibrium is stable if the sensitivity or the reaction time is small. As the sensitivity or the reaction time increases, the equilibrium becomes unstable and the speeds and spacings oscillate. When a forcing term is introduced, most steady-state solutions are periodic, though one of them seem to be quasi-periodic. When the sensitivity or reaction time is small, small disturbances in the first car are damped along the cars and stability over cars is assured. Otherwise, the disturbances are amplified and rear-end collisions may occur as a result.

No evidence of chaos has been found in the car-following model, although other authors claimed to have found chaos solely by calculating a positive Liapunov exponent of a solution of the model (Disbro & Frame, 1990). Chaos should be indicated first of all by irregular motions on a chaotic attractor; Liapunov exponents can only show how fast neighbouring trajectories diverge exponentially and cannot be used to identify chaos on its own. Here, in the numerical analysis in Chapter 4, only equilibria and periodic solutions as attractors have been found in the autonomous system. Introducing a forcing term when the autonomous system has a periodic solution may complicate the behaviour. However, instability over cars occurs for values of parameters which may be smaller than that for the autonomous model to have periodic solutions. When instability over cars occurs, cars collide with each other and the system collapses. It is strongly believed, from the experience of the investigations in this thesis, that chaos is unlikely to occur in the car-following model.

Unfortunately, it has not been possible to prove theoretically that there is definitely no chaos in the car-following model. Generally speaking, if chaos (or any other kind of behaviour) is a typical kind of behaviour in a model, then it can normally be found by numerical calculations. Otherwise, if chaos exists only for some particular values of parameters, then the chances for it to be identified numerically are very small. In addition, it is generally very difficult to prove

that a dynamic model does not have chaos. From the practical point of view, on the other hand, it is probably more important to know when the equilibrium is stable than to show that the model does not possess chaos; the latter is only of theoretical interest, especially when chaos is not a typical kind of behaviour in the model.

7.1.2. The gravity model

The gravity model is a system of difference equations (iterated maps) modelling the variations of the number of trips between each O–D pair at discrete time periods, given the number of trips from each origin and to each destination. The number of trips between each O–D pair at current time period is determined from the travel cost at the previous time period; the travel costs are assumed to increase with the number of trips. Thus there is an interaction between the costs and the flows. It is this interaction which dominates the long term distribution of trips in an O–D network.

The existence, the uniqueness, and the stability of an equilibrium in the model were examined theoretically. Numerical iterations were made under various conditions to find other types of attractors.

With the power or the exponential deterrence function, the distribution of O–D flows approaches a unique equilibrium if the interactions are mild, that is, if trip makers are not too sensitive to changes of travel costs or if the dependence of costs on the flows is not too strong. Otherwise, the equilibrium is unstable and the O–D flows oscillate. The unique equilibrium in the gravity model with the exponential deterrence function gives the stochastic user equilibrium in an O–D network. At the equilibrium, no trip maker can reduce his or her perceived travel cost by changing origins and/or destinations.

When the combined deterrence function is used, changes of the values of a parameter can result in very complicated bifurcation sequences. Irregular behaviour is typical rather than exceptional. Period doubling bifurcations and chaos were found in the model. Sensitive dependence on initial conditions was measured by (positive) Liapunov exponents. In the chaotic regime, the final variations in three or four dimensional systems were found to settle down in the phase space to a region of dimension between 1 and 2, which is neither a one

dimensional line, nor a two dimensional surface.

7.1.3. The logit-based trip assignment model

The logit-based trip assignment model describes the adjustments of flow patterns through a road network at discrete times for a given O–D matrix. The current route flows are determined from the travel costs at the previous time period by the logit choice model; the costs depends on the flows though the link performance functions. Here again, there is an interplay between the flows and costs, like there is in the gravity model. This model was investigated in the same way as that used for the gravity model.

There is a unique equilibrium in the assignment model. The equilibrium is identical to the stochastic user equilibrium. The equilibrium is stable when the inter-dependence between the costs and the flows is not too strong, indicated by smaller values of parameters. It becomes unstable and route flows oscillate as the values of parameters increase.

The dominant forms of behaviour in this model are the stable equilibrium and oscillations. No evidence of more complicated behaviour or chaos has been found in the model.

7.1.4. General conclusions

In Chapter 1 and Chapter 3, we have seen that a very simple dynamic model such as the logistic equation can give rise to very rich dynamic behaviour. Equally, a very complicated model can exhibit rather simple behaviour such as a stable equilibrium. The complexity of the dynamic behaviour of a system is not necessarily proportional to that of the equation of the model. Here, the car-following model and the logit trip assignment model are complicated, but they behave in a simple way. However, the dynamic gravity model does possess a variety of behaviour, from equilibrium to complicated bifurcations and chaos. Chaos is fascinating to analysts, but it is definitely undesirable in practice. Identifying chaos in a traffic model is essential in understanding and avoiding chaos in practice.

The values of parameters for different kind of dynamic behaviour in the three traffic models were given in relevant chapters. What behaviour the system would exhibit depends on ranges of values these parameters might take in practice. For each particular system, only when we know the actual values of parameters in practice, can we know the behaviour of the system through model analyses.

There have been only a limited number of empirical studies of the car-following model (Chandler *et al.*, 1958, Gazis *et al.*, 1961). Although the authors concluded that a nonlinear car-following model is necessary to account for observed traffic flow behaviour (Gazis *et al.*, 1961), no particular model was found to be better than all others. It appears that much more is known about the car-following model theoretically than practically. To test the car-following model, the histories of the speed and the position of each car need to be recorded, from which relative quantities can be obtained. This kind of data is difficult to collect in natural traffic flows.

There has been virtually no empirical study of the dynamic gravity model. The values of parameters in the gravity model are normally different for different areas. Some typical values of parameters in the logit trip assignment models were tried and both stable equilibrium and oscillatory behaviour were found. Traffic data related to trip distribution models and trip assignment models is even more difficult to collect. This kind of data is normally obtained from an O–D survey which is a large scale traffic survey and involves asking each trip maker about their origins and destinations. Clearly, empirical investigations can be more complicated if dynamic considerations are involved.

Both the gravity model with the exponential deterrence function and the trip assignment model considered here have the stochastic user equilibrium as a unique fixed point; this is an attractive property. Sufficient conditions for the equilibria to be stable were provided here. A condition that is both sufficient and necessary is desirable, but it has not been possible to derive such a condition. The difficulties for doing this were explained in the relevant chapters. In addition, the effects of the values of parameters on the stability of the equilibrium in the two models were examined by numerical analysis.

Traffic systems are highly variable and irregular because human behaviour is involved. Therefore we may expect to see more chaos than steadiness in traffic models. However, it is important to remember that traffic models are not real

systems after all. They can only be approximate representations of the system. The variations of traffic flow can be considered to consist of two parts, deterministic variations of average values and random noise. The behaviour found in deterministic models should be considered to explain the variations of average values of the system, or the systematic variations of the system, rather than the random fluctuations. However, in a chaotic regime, it is almost impossible to distinguish deterministic chaos from random noise, as has been mentioned in Chapter 1.

7.2. POSSIBLE EXTENSIONS OF THE RESEARCH

The research in this thesis has been based on theoretical models of traffic flow, selected from the limited number of dynamic traffic models available. It has concentrated on model analysis rather than model development, calibration and validation; the latter three parts are in a much wider area of research. In this section, some possible extensions of the research in this thesis are discussed. The potential of dynamic study based on traffic data is also introduced.

7.2.1. A better equilibrium in the car-following model

We have seen that the equilibrium in the car-following model is not unique; there is a continuum of equilibria. In the car-following model, drivers respond only to the relative speed. Whenever the relative speed becomes zero, the relative spacing will remain constant. The spacing between different cars in the line at an equilibrium is generally different. What is more, these spacings may be very big or very small, which makes traffic either inefficient or dangerous. A further study could be to develop a model that gives an equilibrium where the speed of all cars are equal and the spacing between the cars are equal, too, at some desirable spacing. A favourable driving condition may be achieved when such a car-following model is used in computer-aided driving.

One possible way to derive such a model is to introduce to the car-following model a second term of the form

$$g\,[x_{n-1}(t-\tau) - x_n(t-\tau) - d], \quad n = 2, 3, \cdots, N,$$

where g is an appropriate function, $x_n(t)$ is the position of nth car in the line at time t, d is the desired spacing, and τ is the reaction time. Then a new car-following model can have the form

$$\ddot{x}_n(t) = f[\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)] + g[x_{n-1}(t-\tau) - x_n(t-\tau) - d],$$

$$n = 2, 3, \cdots, N.$$

The first term is a function of relative speed. By choosing suitable forms of f and g a model may be constructed such that it has an equilibrium at which the relative speed is zero and spacing d. In fact, Chandler *et al.* (1958) considered the linear form of this model. However, the second term was found insignificant, using limited car-following data, and was therefore discarded. Further research is needed to find out whether any nonlinear form of the model would be more suitable.

7.2.2. Combined trip distribution, assignment, and modal choice

The gravity trip distribution model determines the O–D flows without specifying which route these flows will go through. In the trip assignment model, the route or link flows are determined for a given trip matrix, which in practice may well be variable. In transport studies, these two models are often used consecutively with the output of trip distribution being used as the input of trip assignment. Before trip assignment is performed the travel costs are unknown and have to be assumed. After the assignment is made the travel costs between O–D pairs are generally different from the ones assumed for trip distribution, since the costs depend on the flows. What is more, the travel cost between an O–D pair in the trip distribution model is rather a vague concept because it is not associated with any particular route. One way round all this is to combine the two processes into one model. The combination may be extended to include modal choice as well, so that the choice of routes, origins or destinations, and travel modes can be made simultaneously. There have been some studies in this direction (Evans, 1976, Florian & Nguyen, 1978, Erlander, 1990, Lam & Huang, 1992), although almost all of them are made in a static context. The kind of dynamic considerations used in this thesis may be used in the combined trip distribution, assignment, and modal choice. Traffic models in this context will be much more complicated since they involve two or three dimensions of choice.

7.2.3. Empirical dynamics in traffic data

As an alternative to a theoretical model, a dynamical system may be described by observed data of one or more state variables of the system, normally in the form of a time series. A phase-space picture, or an attractor, may be reconstructed from a time series of one single variable of the system (Packard *et al.* 1980). *Empirical chaos*, or chaos in observed data, can be detected (if it exists) by calculating Liapunov exponents (Wolf *et al.*, 1985, Eckmann *et al.*, 1986, Conte & Dubois, 1988) and fractal dimensions (Grassberger & Procaccia, 1983) from the reconstructed attractor. No existing study has been made on empirical chaos in traffic systems, but empirical chaos has been found in some other dynamical systems, such as in economic systems and in chemical reaction systems (Roux *et al.*, 1983, Swinney, 1983, Frank & Stengos, 1988).

However, to carry out this kind of investigation, the time series needs to be long, measured with adequate precision and with little random noise. A suitable time series for study might contain 40,000 points with a precision of four effective numbers. Such kind of data often has to be collected in carefully controlled experiments in order to reduce noise. The collection of traffic data is inevitably subject to disturbing factors such as traffic incidents, weather conditions, and so on. Some of the factors are not predictable and so cannot be controlled. Therefore, traffic data collected under natural conditions can contain very large stochastic components. Dynamic analysis based on this kind of data may be meaningless and even misleading. However, there is a potential area for obtaining this kind of data, namely, the car-following process. Car-following experiments may be made under controlled conditions. In fact, some early empirical studies of car-following mentioned in this thesis were made under controlled conditions (Chandler *et al.*, 1958, Gazis *et al.*, 1961). However, the data collected is far too short for investigating empirical dynamics in the car-following process. With the development of car-following simulators it may be possible in the near future to obtain the data needed. It will then be possible to test for the presence of empirical chaos in the car-following process.

As a side product of this thesis, programs for calculating Liapunov exponents and fractal dimensions from time series data have been developed based on the algorithms by Eckmann *et al.* (1986) and Grassberger & Procaccia (1983). The

programs have been used successfully by two undergraduate students (Peters, 1993, Unlu, 1995) in Middlesex University to analyze time series data of economic systems. See Appendix B for listings of the programs.

7.3. A final comment

To conclude, let us go back to the title of this thesis — *The dynamic behaviour of road traffic flow: stability of chaos?* Traffic is often chaotic in the ordinary sense. Whether or not it is chaotic in terms of nonlinear dynamics may be a different matter altogether. Stability and equilibrium are desirable; instability and chaos in any way are to be avoided. In order to achieve the desired traffic state, we need first of all to understand and to predict traffic behaviour.

In this thesis, the dynamic behaviour of traffic flows on a road link, in a road network, and in an O–D network of an area has been identified for various conditions. A future step could be to modify the performance of traffic systems by better dynamic modelling, based on sound empirical analysis, and probably complemented by optimum control methodology. With these perfected models, we may then be able to monitor and control real traffic systems so that they operate in the desired state.

APPENDIX A. THE ALGORITHM FOR INTEGRATING THE CAR-FOLLOWING EQUATIONS

In this appendix, the algorithm used for integrating the car-following equations in Chapter four is described. The car-following model is a set of delay-differential equations. The fourth order Runge-Kutta method for solving ordinary differential equations is modified so as to deal with delay-differential equations. The modification used here is the same to Farmer's (Farmer, 1982), but the algorithm is adapted so as to integrate the system of car-following equations. The Runge-Kutta method is described first and then modifications to it follow.

Consider an initial value problem of a system of two ordinary differential equations

$$\begin{aligned} \dot{X} &= P(t,X,Y), \\ \dot{Y} &= Q(t,X,Y), \\ X(t_0) &= X_0, \\ Y(t_0) &= Y_0. \end{aligned}$$

In numerical integrations the time t is normally discretized as

$$t(i+1) = t(i) + h, \quad i = 0, 1, \cdots,$$

where h is the step length. The solution at time $t(i+1)$ by Runge-Kutta algorithm is (Conte and Boor, 1980)

$$\begin{aligned} X(i+1) &= X(i) + \frac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4), \\ Y(i+1) &= Y(i) + \frac{1}{6}(l_1 + 2l_2 + 2l_3 + l_4), \\ i &= 0, 1, \cdots, \end{aligned}$$

where

$$\begin{aligned} k_1 &= h\,P[t(i),\ X(i),\ Y(i)], \\ l_1 &= h\,Q[t(i),\ X(i),\ Y(i)], \end{aligned}$$

$$
\begin{aligned}
k_2 &= h\,P[t(i) + h/2,\ X(i) + k_1/2,\ Y(i) + l_1/2],\\
l_2 &= h\,Q[t(i) + h/2,\ X(i) + k_1/2,\ Y(i) + l_1/2],\\
k_3 &= h\,P[t(i) + h/2,\ X(i) + k_2/2,\ Y(i) + l_2/2],\\
l_3 &= h\,Q[t(i) + h/2,\ X(i) + k_2/2,\ Y(i) + l_2/2],\\
k_4 &= h\,P[t(i) + h,\ X(i) + k_3,\ Y(i) + l_3],\\
l_4 &= h\,Q[t(i) + h,\ X(i) + k_3,\ Y(i) + l_3].
\end{aligned}
$$

The standard Runge-Kutta method for integrating ordinary differential equations can be extended to integrate delay-differential equations. We will illustrate the extension by an example first. Suppose the delay-differential equation to be integrated is

$$\dot{\mathbf{Z}} = \mathbf{H}(t, \mathbf{Z}, \mathbf{Z}(t-\tau)),$$

where

$$
\begin{aligned}
\mathbf{Z} &= (Z_1, Z_2, \ldots, Z_M)^{\mathrm{T}},\\
\mathbf{Z}(t-\tau) &= (Z_1(t-\tau), Z_2(t-\tau), \ldots, Z_M(t-\tau))^{\mathrm{T}},
\end{aligned}
$$

and the superscripts "T" denote transpose. In this kind of differential equation the evolution is determined by the function $\mathbf{Z}$ on the time interval $[t, t-\tau]$. This function can be approximated by T points taken at intervals $h = \tau/(T-1)$. Thus, $\mathbf{Z}(t)$ and $\mathbf{Z}(t-\tau)$ can be expressed as $\mathbf{Z}(i)$ and $\mathbf{Z}(i-T)$ respectively when discretized in the Runge-Kutta formulae, as follows

$$Z_n(i+1) = Z_n(i) + \frac{1}{6}\,(k_n^1(i) + 2k_n^2(i) + 2k_n^3(i) + k_n^4(i)), \qquad \text{(A.1)}$$

$$n = 1, 2, \cdots, M, \quad i = 0, 1, \cdots,$$

where

$$
\begin{aligned}
k_n^1(i) &= h\,H_n[t(i),\ \mathbf{Z}(i),\ \mathbf{Z}(i-T)],\\
k_n^2(i) &= h\,H_n[t(i) + h/2,\ \mathbf{Z}(i) + \boldsymbol{k}^1(i)/2,\ \mathbf{Z}(i-T) + \boldsymbol{k}^1(i-T)/2],\\
k_n^3(i) &= h\,H_n[t(i) + h/2,\ \mathbf{Z}(i) + \boldsymbol{k}^2(i)/2,\ \mathbf{Z}(i-T) + \boldsymbol{k}^2(i-T)/2],\\
k_n^4(i) &= h\,H_n[t(i) + h,\ \mathbf{Z}(i) + \boldsymbol{k}^3(i),\ \mathbf{Z}(i-T) + \boldsymbol{k}^3(i-T)].
\end{aligned}
$$

Here, the superscripts are notations and do not stand for powers. This notation will be used in all Runge-Kutta formulae throughout this appendix. The vectors in the above formulae are

$$\mathbf{Z}(i) = (Z_1(i), Z_2(i), ..., Z_M(i))^{\mathrm{T}},$$
$$\mathbf{Z}(i-T) = (Z_1(i-T), Z_2(i-T), ..., Z_M(i-T))^{\mathrm{T}},$$
$$\boldsymbol{k}^j(i) = (k_1^j(i), k_2^j(i), \cdots, k_n^j(i))^{\mathrm{T}},$$
$$\boldsymbol{k}^j(i-T) = (k_1^j(i-T), k_2^j(i-T), \cdots, k_n^j(i-T))^{\mathrm{T}},$$

for $j = 1, 2, 3, 4,$ and $i = 0, 1, \cdots.$

To simplify the expression of the formulae for $k_n^j(i)$'s, let

$$Z_n^1(i) \equiv Z_n(i),$$
$$Z_n^2(i) \equiv Z_n(i) + k_n^1(i)/2,$$
$$Z_n^3(i) \equiv Z_n(i) + k_n^2(i)/2,$$
$$Z_n^4(i) \equiv Z_n(i) + k_n^3(i),$$
$$t^1(i) \equiv t(i),$$
$$t^2(i) \equiv t(i) + h/2,$$
$$t^3(i) \equiv t(i) + h/2,$$
$$t^4(i) = t(i) + h.$$

Then the formulae for $k_n^j(i)$'s in (A.1) become

$$k_n^j(i) = h\, H_n[t^j(i), \mathbf{Z}^j(i), \mathbf{Z}^j(i-T)], \qquad (A.2)$$
$$j = 2, 3, 4, \quad n = 1, 2, \cdots, M, \quad \text{and} \quad i = 0, 1, \cdots.$$

The initial conditions are given by specifying $\mathbf{Z}(i)$ and $\boldsymbol{k}^j(i)$ for $i = -T, -T+1, \cdots, 0,$ and $j = 1, 2, 3, 4.$ Given an initial condition, the integration can be proceeded as follows. For each $i,$ calculate $k_n^j(i)$ for $n = 1, 2, \cdots, M,$ and $j = 1, 2, 3, 4,$ from which $\mathbf{Z}(i+1)$ can be obtained. Repeat the process until the desired integration time is reached.

In the standard Runge-Kutta method for integrating ordinary differential equations, only the solution at the ith step is needed for calculating the solution at the $(i+1)$th step. In the case of delay-differential equations, however, the

the values of $\mathbf{Z}^j(I)$ for j=1, 2, 3, 4, and for $I = i, i-1, i-2, \cdots, i-T$ have to be saved for further calculations, which works out an array of the size $M \times T \times 4$.

Now consider the car-following model (4.2) in Chapter 4

$$\dot{v}_2(t) = \ddot{x}_1(t) - \beta_2(t)\, v_2(t-\tau),$$
$$\dot{v}_n(t) = \beta_{n-1}(t)\, v_{n-1}(t-\tau) - \beta_{n-1}(t)\, v_{n-1}(t-\tau), \quad n = 3, 4, \ldots, N,$$
$$\dot{y}_n(t) = v_n(t), \quad n = 2, 3, \ldots, N,$$

where

$$\beta_n(t) = \frac{\alpha\ (\dot{x}_1(t) - v_2(t) - \cdots - v_n(t))^m}{(y_n(t-\tau) + b)^l}, \quad n = 2, 3, \ldots, N.$$

The car-following model is a set of recursive equations; each equation contains only the variables of two adjacent cars. Therefore, there will be lots of unnecessary repeating calculations if the modified Runge-Kutta formulae (A.1) and (A.2) are applied directly to integrate the model. This can be avoided by the following changes.

Let

$$u_n(t) = \dot{x}_n(t), \qquad n = 1, 2, \cdots, N.$$

Then

$$v_n(t) = \dot{x}_{n-1}(t) - \dot{x}_n(t) = u_{n-1}(t) - u_n(t), \quad n = 1, 2, \cdots, N,$$

and so

$$u_{n-1} = \dot{x}_1(t) - v_2(t) - \cdots - v_{n-1}(t), \quad n = 3, 4, \cdots, N.$$

The car-following model can now be written as

$$\dot{v}_n(t) = \ddot{x}_{n-1}(t) - \ddot{x}_n(t), \quad n = 2, 3, \ldots, N,$$
$$\dot{y}_n(t) = v_n(t), \qquad n = 2, 3, \ldots, N.$$

where

$$\ddot{x}_n(t) = \frac{\alpha(u_{n-1}(t) - v_n(t))^m}{(y_n(t-\tau) + b)^l} v_n(t-\tau) \equiv f_n(u_{n-1}(t), v_n(t), v_n(t-\tau), y_n(t-\tau)),$$

$$n = 2, 3, \ldots, N.$$

Here, we have used f_n to denote the acceleration of car n

$$f_n(u_{n-1}(t),\ v_n(t),\ v_n(t-\tau),\ y_n(t-\tau)) = \frac{\alpha(u_{n-1}(t) - v_n(t))^m}{(y_n(t-\tau) + b)^l} v_n(t-\tau) \tag{A.3}$$

When the formulae (A.1) and (A.2) are applied to the above equations, we obtain

$$v_n(i+1) = v_n(i) + \frac{1}{6}(k_n^1(i) + 2k_n^2(i) + 2k_n^3(i) + k_n^4(i)),$$
$$y_n(i+1) = y_n(i) + \frac{1}{6}(l_n^1(i) + 2l_n^2(i) + 2l_n^3(i) + l_n^4(i)),$$

$$n = 2, 3, \cdots, N, \text{ and } i = 0, 1, \cdots,$$

where

$$k_n^j(i) = h\ (\ddot{x}_{n-1}^j(i) - \ddot{x}_n^j(i)) \quad j = 1, 2, 3, 4,$$
$$l_n^j(i) = h\ v_n^j(i) \quad j = 1, 2, 3, 4.$$

By (A.3), we have

$$\ddot{x}_n^1(i) = f_n(u_{n-1}^1(i),\ v_n(i),\ v_n(i-T),\ y_n(i-T)),$$
$$\ddot{x}_n^2(i) = f_n(u_{n-1}^2(i),\ v_n(i)+k_n^1(i)/2,\ v_n(i-T)+k_n^1(i-T)/2,$$
$$y_n(i-T)+l_n^1(i-T)/2),$$
$$\ddot{x}_n^3(i) = f_n(u_{n-1}^3(i),\ v_n(i)+k_n^2(i)/2,\ v_n(i-T)+k_n^2(i-T)/2,$$
$$y_n(i-T)+l_n^2(i-T)/2),$$
$$\ddot{x}_n^4(i) = f_n(u_{n-1}^4(i),\ v_n(i)+k_n^3(i),\ v_n(i-T)+k_n^3(i-T),$$
$$y_n(i-T)+l_n^3(i-T)),$$

or simply,

$$\ddot{x}_n^j(i) = f_n(u_{n-1}^j(i),\ v_n^j(i),\ v_n^j(i-T),\ y_n^j(i-T)),$$

where

$$u_n^j(i) = u_{n-1}^j(i) - v_n^j(i),$$

$$v_n^1(i) = v_n(i),$$
$$v_n^2(i) = v_n(i)+k_n^1(i)/2,$$
$$v_n^3(i) = v_n(i)+k_n^2(i)/2,$$
$$v_n^4(i) = v_n(i)+k_n^3(i),$$

$$y_n^1(i) = y_n(i),$$
$$y_n^2(i) = y_n(i)+l_n^1(i)/2,$$
$$y_n^3(i) = y_n(i)+l_n^2(i)/2,$$
$$y_n^4(i) = y_n(i)+l_n^3(i).$$

Defined by the motion of the first car we have

$$\ddot{x}_1^1(i) = \ddot{x}_1(t(i)),$$
$$\ddot{x}_1^2(i) = \ddot{x}_1^3(i) = \ddot{x}_1(t(i)+h/2),$$
$$\ddot{x}_1^4(i) = \ddot{x}_1(t(i)+h),$$

and

$$u_1^1(i) = \dot{x}_1(t(i)),$$
$$u_1^2(i) = u_1^3(i) = \dot{x}_1(t(i)+h/2),$$
$$u_1^4(i) = \dot{x}_1(t(i)+h),$$

for i=1, 2, $\cdots$.

The initial functions are given by specifying $v_n(i)$, $k_n^j(i)$, $y_n(i)$, and $l_n^j(i)$ for $n = 1, 2, \cdots, N$, $j = 1, 2, 3, 4$, and $i = -T, -T+1, \cdots, 0$.

In this algorithm, $\ddot{x}_n^j(i)$ and $u_n^j(i)$ are calculated and stored explicitly so that they can be saved for calculating $k_{n+1}^j(i)$ and so $v_{n+1}(i+1)$ to avoid repetitive calculations. In this way, computer time can be reduced.

APPENDIX B. LISTINGS OF SOURCE PROGRAMS

The programs used in this thesis are all written in FORTRAN. Six major programs are listed in this appendix. They are:

(1) program CARFL, which integrates the car-following equations;
(2) program UNCLE, which calculates Liapunov exponents of an attractor of the unconstrained or singly constrained gravity model;
(3) program DBCLE, which calculates Liapunov exponents of an attractor of the doubly constrained gravity model;
(4) program CORDIM, which calculates the correlation function of an attractor;
(5) program TSLE, which calculates Liapunov exponents from time series data of a single variable of a dynamical system;
(6) program TSDIM, which calculates the correlation function of an attractor reconstructed from time series data of a single variable of a dynamical system.

The algorithm, the input and the output data used in each program are introduced in that program. The method for calculating Liapunov exponents of an attractor of unconstrained or singly constrained gravity model is significantly different from that for the doubly constrained gravity model, as we have seen in Chapter 5. Therefore, the programs are coded separately. Programs CORDIM and TSDIM produces the logarithms of both the distances and the values of correlation functions. These are ready to be used to obtain correlation dimension of an attractor by, for example, MATLAB (The MathWorks Inc, 1993). The method for doing this has been described in section 5.5.1.

It has been mentioned in Chapter 7 that a dynamical system may be described by observed data, normally in the form of time series, of the system. Liapunov exponents and correlation dimension can be calculated from time series data of one single variable of the system (Packard *et al.* 1980, Grassberger and Procaccia, 1983, Wolf *et al.*, 1985, Eckmann *et al.*, 1986, and Conte and Dubois, 1988) to

detect possible empirical chaos. Programs TSLE and TSDIM calculate Liapunov exponents and correlation function, respectively, of an attractor of such dynamical systems. The algorithm used in TSLE is due to Eckmann *et al.* (1986) and the algorithm used in TSDIM due to Grassberger and Procaccia (1983).

The programs CARFL, UNCLE, DBCLE, and CORDIM have been tested and used in the relevant chapters in this thesis. The programs TSLE and TSDIM were tested by using the time series data of one single variable of the Hénon map mentioned in the thesis. Liapunov exponents and correlation dimension of a chaotic attractor in the map were calculated by programs TSLE and TSDIM respectively. The results were compared with the known values in the literature (Conte and Dubois, 1988, and Grassberger and Procaccia, 1983) and good agreements were found. These two programs have been used successfully by two undergraduate students (Peters, 1993, Unlu, 1995) in Middlesex University to analyze time series data of economic systems.

B.1. THE PROGRAM CARFL

```
      PROGRAM CARFL
C
C This program integrates the delay-differential equation of the
C car-following model by modified fourth order Runge-Kutta
C algorithm.  The algorithm is described in Appendix A.
C
C The input data:
C KV    = number of cars considered;
C REACT = reaction time;
C SENS  = coefficient of proportionality;
C M     = parameter in the car-following equation;
C L     = parameter in the car-following equation;
C Y0    = minimum headway;
C A     = amplitude of the forcing term;
C B     = average speed of the first car;
C C     = frequency of the forcing term;
C X0(I) = initial value of relative spacing of car I+1 to car I;
C V0(I) = initial value of relative speed of car I+1 to car I.
C The input data is in the data file CARFL1.DAT.
C
C The output data:
C YL(I) = relative spacing of car I+1 to car I;
C VL(I) = relative speed of car I+1 to car I.
C There are two output data files:
C CARFL2.DAT contains the transient solutions;
C CARFL3.DAT contains the steady state solutions.
C
      PARAMETER (NCAR=10,NPT=2000)
      DOUBLE PRECISION V(NCAR,NPT),Y(NCAR,NPT),
     1  YL(NCAR),YN(NCAR),VL(NCAR),VN(NCAR),ALE(4),VLE(4),
     2  AL1(NCAR,NPT),AL2(NCAR,NPT),AL3(NCAR,NPT)
      REAL M,L,X0(NCAR),V0(NCAR),H,TL,
     1  VIN1,VIN2,TIME,REACT,SENS,Y0,A,B,C
      INTEGER KV,MT,ITW,BWT,NT
      OPEN(UNIT=11,FILE='CARFL1.DAT',STATUS='OLD',
     1  FORM='FORMATTED',ACCESS='SEQUENTIAL')
      OPEN(UNIT=12,FILE='CARFL2.DAT',STATUS='UNKNOWN',
     1  FORM='FORMATTED',ACCESS='SEQUENTIAL')
      OPEN(UNIT=13,FILE='CARFL3.DAT',STATUS='UNKNOWN',
     1  FORM='FORMATTED',ACCESS='SEQUENTIAL')
      READ(11,*)KV,H,REACT,MT,TIME,ITW,BWT,
     1  Y0,A,B,C,M,L,SENS,VIN1,VIN2
      READ(11,*) (X0(I),I=1,KV-1),(V0(I),I=1,KV-1)
C
C Initialize
C
      DO 30 K=1,KV-1
        TIN=0
        DO 20 I=1,MT
          V(K,I)=FINV(V0(K),TIN,VIN1,VIN2)
          AL1(K,I)=FINV(V0(K),TIN+H/2,VIN1,VIN2)
          AL2(K,I)=AL1(K,I)
          AL3(K,I)=FINV(V0(K),TIN+H,VIN1,VIN2)
          Y(K,I)=FINY(V0(K),X0(K),TIN,VIN1,VIN2)
          TIN=TIN+H
 20     CONTINUE
        VL(K)=FINV(V0(K),MT*H,VIN1,VIN2)
        YL(K)=FINY(V0(K),X0(K),MT*H,VIN1,VIN2)
        YN(K)=0.0
        VN(K)=0.0
 30   CONTINUE
```

```
      TL=0
      NT=TIME/REACT
C
C Start integrating
C
      DO 200 I=1,NT
        DO 100 J=1,MT
C Calculate the speed and the acceleration of the first car
          CALL LEAD(TL,A,B,C,ALE(1),VLE(1))
          CALL LEAD(TL+H/2,A,B,C,ALE(2),VLE(2))
          ALE(3)=ALE(2)
          VLE(3)=VLE(2)
          CALL LEAD(TL+H,A,B,C,ALE(4),VLE(4))
C Calculate the relative speed and spacing among following cars
          DO 50 K=1,KV-1
            CALL AKTY(TL,H,V(K,J),VL(K),Y(K,J),YL(K),AL1(K,J),
     1      AL2(K,J),AL3(K,J),VN(K),YN(K),ALE,VLE,Y0,M,L,SENS)
            V(K,J)=VL(K)
            VL(K)=VN(K)
            Y(K,J)=YL(K)
            YL(K)=YN(K)
 50       CONTINUE
          TL=TL+H
C Print the current solution
          JJ=((I-1)*MT+J)/ITW
          RJ=FLOAT((I-1)*MT+J)/FLOAT(ITW)
          IF(JJ.LT.RJ) GOTO 100
          IF(TL.LT.BWT) THEN
            WRITE(12,'(1X,7F10.4)') TL,(YL(II),II=1,KV-1),
     1      (VL(II),II=1,KV-1)
            WRITE(*,'(1X,7F10.4)') TL,(YL(II),II=1,KV-1),
     1      (VL(II),II=1,KV-1)
          ELSE
            WRITE(13,'(1X,7F10.4)') TL,(YL(II),II=1,KV-1),
     1      (VL(II),II=1,KV-1)
            WRITE(*,'(1X,7F10.4)') TL,(YL(II),II=1,KV-1),
     1      (VL(II),II=1,KV-1)
          ENDIF
 100    CONTINUE
 200  CONTINUE
      STOP
      END

      SUBROUTINE AKTY(T,H,VJ,VL,YJ,YL,AL1,AL2,AL3,VN,YN,
     1  ALE,VLE,Y0,M,L,SENS)
C
C This subroutine integrates the car-following equation
C for one step by modified fourth order Runge-Kutta formula.
C
      REAL T,H,M,L,Y0,SENS
      DOUBLE PRECISION VJ,VL,VN,YJ,YL,YN,DY,DV,
     1  ALE(4),VLE(4),RA(4),RV(4),AL1,AL2,AL3,
     1  AK1,AK2,AK3
      RV(1)=VL
      RA(1)=CFE(ALE(1),VLE(1),RV(1),VJ,YJ,Y0,M,L,SENS)
      AK1=H*RA(1)
      RV(2)=VL+AK1/2
      RA(2)=CFE(ALE(2),VLE(2),RV(2),AL1,YJ+H*VJ/2,Y0,M,L,SENS)
      AK2=H*RA(2)
      RV(3)=VL+AK2/2
      RA(3)=CFE(ALE(3),VLE(3),RV(3),AL2,YJ+H*AL1/2,Y0,M,L,SENS)
      AK3=H*RA(3)
      RV(4)=VL+AK3
      RA(4)=CFE(ALE(4),VLE(4),RV(4),AL3,YJ+H*AL2,Y0,M,L,SENS)
      AK4=H*RA(4)
      DV=(AK1+2*AK2+2*AK3+AK4)/6
      VN=VL+DV
      DY=H*VL+H*(AK1+AK2+AK3)/6
```

```
      YN=YL+DY
      DO 1000 LL=1,4
        ALE(LL)=ALE(LL)-RA(LL)
        VLE(LL)=VLE(LL)-RV(LL)
 1000 CONTINUE
      AL1=RV(2)
      AL2=RV(3)
      AL3=RV(4)
      RETURN
      END

      DOUBLE PRECISION FUNCTION CFE(ALC,VLC,VL,VJ,YJ,Y0,M,L,SENS)
C
C Calculate the value of the function defined by the
C car-following equation.
C
      REAL M,L,Y0,SENS
      DOUBLE PRECISION VL,VJ,YL,ALC,VLC
      X1=VLC-VL
C Check the negativity of the speed
      IF(X1.LT.0.0) THEN
        PRINT*,'X1=VLC-VL',X1,VLC,VL
        STOP
      ENDIF
C Check the boundness of the spacing
      IF(X1.GT.200.0) THEN
        PRINT*,'X1=VLC-VL',X1,VLC,VL
        STOP
      ENDIF
      X2=YJ+Y0
C Check the negativity of the spacing
      IF(X2.LT.0.0) THEN
        PRINT*,'X2=YJ+Y0',X2,YJ,Y0
        STOP
      ENDIF
C Evaluate the function
      X1=X1**M
      X2=X2**L
      CFE=ALC-SENS*X1*VJ/X2
      RETURN
      END

      SUBROUTINE LEAD(T,A,B,C,ALC,VLC)
C
C Calculate the acceleration and the speed of the first car
C
      DOUBLE PRECISION ALC,VLC
      REAL T,A,B,C
      ALC=A*SIN(C*T)
      VLC=-A/C*COS(C*T)+B
      RETURN
      END

      REAL FUNCTION FINV(V0,T,A,C)
C
C Evaluate the initial function of relative speed
C
      REAL V0,T,A,C
      FINV=V0+A*SIN(C*T)
      RETURN
      END

      REAL FUNCTION FINY(V0,X0,T,A,C)
C
C Evaluate the initial function of relative spacing
C
      REAL V0,X0,T,A,C
      FINY=X0+V0*T+(A/C)*(1-COS(C*T))
```

```
RETURN
END
```

B.2. THE PROGRAM UNCLE

```
      PROGRAM UNCLE
C
C This program calculates Liapunov exponents for the
C attractor of the unconstrained or singly
C constrained gravity model by the algorithm by
C Eckmann and Ruelle (1985).
C The algorithm is described in section 5.4.
C The program uses NAG routines (NAG Ltd, 1987):
C F01CKF, F01QCF, and F01QEF.
C
C The input data:
C NIJ   = number of O-D pair;
C Z0(I) = relative capacity of O-D pair I;
C C0(I) = uncongested travel cost of O-D pair I;
C X0(I) = initial number of trips of O-D pair I;
C ALPHA = parameter in the model;
C BETA  = parameter in the model;
C GAMA  = parameter in the model;
C RN    = parameter in the model.
C Different values of the above parameters may be tried.
C
C The output data:
C X(I)  = solution or the number trips of O-D pair I;
C RLE(I)= the Ith Liapunov exponent, which is contained
C       in the data file UNCLE1.DAT.
C
      PARAMETER (LDIM=3,NIJ=4)
      INTEGER IFAIL
      DOUBLE PRECISION C0(NIJ),Z0(NIJ),ALPHA,GAMA,RN,BETA,
     1  BETA1(5),RN1(5),FC,CY,C(NIJ),F(NIJ),FPRIME(NIJ),RK
      DOUBLE PRECISION X0(NIJ),X(NIJ),XL(NIJ),T(LDIM,LDIM),
     1  Q(LDIM,LDIM),R(LDIM,LDIM),CUMR(LDIM),RLE(LDIM),
     2  WORK(LDIM),Z(LDIM),ZETA(LDIM)
      OPEN(1,FILE='UNCLE1.DAT',STATUS='UNKNOWN',FORM='FORMATTED')
C
C Input data
C
      DATA ALPHA,GAMA/1.0, 1.0/
      DATA (BETA1(I),I=1,1)/2.5/
      DATA (RN1(I),I=1,1)/7.0/
      DATA (Z0(I),I=1,4)/
     1  0.17, 0.15, 0.25, 0.23/
      DATA (C0(I),I=1,4)/
     1  1.40, 1.20, 1.80, 1.60/
      DATA (X0(I),I=1,4)/
     1  0.25, 0.25, 0.25, 0.25/
C
C Iteration with different values of parameters
C
      DO 9500 IIII=1,1
      RN=RN1(IIII)
      DO 9500 JJJJ=1,20
      BETA=BETA1(1)+(JJJJ-1)*0.05
C
C Initialize
C
        DO 100 I=1,NIJ
          X(I)=X0(I)
 100    CONTINUE
        DO 80 I=1,LDIM
          DO 50 J=1,LDIM
```

```
            Q(I,J)=0.0
 50         CONTINUE
            CUMR(I)=0.0
            Q(I,I)=1.0
 80       CONTINUE
          ITER=10000
          IO=5000
          KB=0
C
C Begin the iteration with time
C
          DO 9000 III=1,ITER
C
C Find current number of trips
C
            DO 200 I=1,NIJ
              XL(I)=X(I)
 200        CONTINUE
            RK=0.0
            DO 500 I=1,NIJ
              C(I)=C0(I)*(1+ALPHA*(XL(I)/Z0(I))**GAMA)
              F(I)=C(I)**RN*EXP(-BETA*C(I))
              RK=RK+F(I)
 500        CONTINUE
            RK=1/RK
            DO 700 I=1,NIJ
              X(I)=RK*F(I)
 700        CONTINUE
C
C Check if transient stage has passed
C
            IF(III.LT.KB) GOTO 7500
C
C Find derivatives
C
            DO 2400 I=1,NIJ
              FC=RN*C(I)**(RN-1)*EXP(-BETA*C(I))-
     1           BETA*C(I)**RN*EXP(-BETA*C(I))
              CY=C0(I)*ALPHA*GAMA*(XL(I)/Z0(I))**(GAMA-1)/Z0(I)
              FPRIME(I)=FC*CY
 2400       CONTINUE
C
C Calculate the Jacobian matrix T
C
            DO 2500 I=1,NIJ-1
            DO 2500 J=1,NIJ-1
              IF(I.EQ.J) THEN
                T(I,J)=RK*(FPRIME(I)-RK*F(I)*(FPRIME(I)-
     1          FPRIME(NIJ)))
              ELSE
                T(I,J)=-RK*RK*F(I)*(FPRIME(J)-FPRIME(NIJ))
              ENDIF
 2500       CONTINUE
C
C Multiply matrix T by the previous orthogonal matrix Q
C
            IFAIL=0
            CALL F01CKF(R,T,Q,LDIM,LDIM,LDIM,Z,LDIM,1,IFAIL)
C
C Do QR factorization of matrix T'= T*Q
C
            CALL F01QCF(LDIM,LDIM,R,LDIM,ZETA,IFAIL)
            DO 6300 I=1,LDIM
            DO 6300 J=1,LDIM
              Q(I,J)=R(I,J)
 6300       CONTINUE
            CALL F01QEF('S',LDIM,LDIM,LDIM,Q,LDIM,ZETA,WORK,IFAIL)
C
```

```
C Calculate Liapunov Exponents and print out
C
          DO 7000 I=1,LDIM
            CUMR(I)=CUMR(I)+DLOG(ABS(R(I,I)))
 7000     CONTINUE
 7500     IF(MOD(III,IO).NE.0) GOTO 9000
          IF(III.LT.KB) GOTO 9000
          DO 8000 I=1,LDIM
            RLE(I)=CUMR(I)/III
 8000     CONTINUE
 9000   CONTINUE
        WRITE(*,'(1X,8F9.5)') BETA,(X(I),I=1,NIJ),(RLE(I),I=1,LDIM)
        WRITE(1,'(1X,8F9.5)') BETA,(X(I),I=1,NIJ),(RLE(I),I=1,LDIM)
 9500 CONTINUE
      STOP
      END
```

B.3. THE PROGRAM DBCLE

```
      PROGRAM DBCLE
C
C This program calculates Liapunov exponents for the
C attractor of the doubly constrained gravity model
C by the algorithm by Eckmann and Ruelle (1985).
C The algorithm is described in section 5.4.
C The program uses NAG routines (NAG Ltd, 1987):
C F01CKF, F01QCF, F01QEF, and F04ATF.
C
C The input data:
C NI    = number of origin;
C NJ    = number of destination;
C Z(I)  = relative capacity of O-D pair I;
C C0(I) = uncongested travel cost of O-D pair I;
C O(I)  = total number of trips from origin I;
C D(J)  = total number of trips to destination J;
C X0(I) = initial number of trips of O-D pair I;
C ALPHA = parameter in the model;
C BETA  = parameter in the model;
C GAMA  = parameter in the model;
C RN    = parameter in the model.
C
C The output data:
C X(I)  = solution or the number trips of O-D pair I;
C RLE(I)= the Ith Liapunov exponent, which is contained
C       in the data file DBCLE1.DAT.
C
      PARAMETER (NN=30,NZ=5,LDIM=4)
      DOUBLE PRECISION C0(NN),Z(NN),X0(NN),OI(NZ),DJ(NZ),
     1  XL(NN),X(NN),C(NN),F(NN),
     2  ALPHA1(10),GAMA1(10),BETA1(10),RN1(10),
     3  ALPHA,GAMA,RN,BETA,ERROR,AI(NZ),BJ(NZ),
     4  FC,CX,FPRIME(NN),T(LDIM,LDIM),Q(LDIM,LDIM),
     5  R(LDIM,LDIM),CUMR(LDIM),RLE(LDIM),
     6  WORK(LDIM),ZWK(LDIM),ZETA(LDIM)
      INTEGER NI,NJ,NIJ,ITER,KB,IFAIL
      OPEN(1,FILE='DBCLE1.DAT',STATUS='UNKNOWN',FORM='FORMATTED')
C Input data
      DATA ERROR/0.00005/
      DATA NI,NJ,NIJ/3,3,9/
      DATA (ALPHA1(I),I=1,1)/1.5/,(GAMA1(I),I=1,1)/1.5/
      DATA (BETA1(I),I=1,1)/1.25/,(RN1(I),I=1,1)/4.50/
      DATA (Z(I),I=1,9)/
     1          0.10,0.11,0.07,
     2          0.09,0.09,0.11,
     3          0.10,0.09,0.07/
      DATA (C0(I),I=1,9)/
     1          1.,1.,1.5,
     2          1.2,1.4,1.8,
     3          1.5,0.9,0.5/
      DATA (X0(I),I=1,9)/
     1          0.050, 0.050, 0.200,
     2          0.100, 0.150, 0.050,
     3          0.200, 0.150, 0.050/
      DATA (OI(I),I=1,3)/0.35,0.35,0.3/
      DATA (DJ(J),J=1,4)/0.3,0.3,0.4,1.0/
      NIJ=NI*NJ
      ITER=5000
      KB=0
      IO=10
C
```

```
C Iteration with different values of parameters
C
      DO 9000 LLLL=1,1
      ALPHA=ALPHA1(LLLL)
      DO 9000 MMMM=1,1
      GAMA=GAMA1(MMMM)
      DO 9000 IIII=1,1
      RN=RN1(IIII)
      DO 9000 JJJJ=1,1
        BETA=BETA1(JJJJ)
        WRITE(*,'(1X,18Halpha,gama,rn,beta,4F10.4)')
     1  alpha,gama,rn,beta
        DO 20 I=1,NIJ
          X(I)=X0(I)
 20     CONTINUE
        WRITE(*,'(1X,9HInitially,9F7.3)') (X(I),I=1,NIJ)
C
C Initialize
C
        DO 80 I=1,LDIM
          DO 50 J=1,LDIM
            Q(I,J)=0.0
 50       CONTINUE
          CUMR(I)=0.0
          Q(I,I)=1.0
 80     CONTINUE
C
C Begin the iteration with time
C
        DO 5000 KK=1,ITER
C
C Find the current cost, the value of deterrence function,
C and the derivatives of the deterrence function with
C respect of number of trips
C
          DO 500 I=1,NIJ
            XL(I)=X(I)
            C(I)=C0(I)*(1+ALPHA*(XL(I)/Z(I))**GAMA)
            F(I)=C(I)**RN*EXP(-BETA*C(I))
            FC=RN*C(I)**(RN-1)*EXP(-BETA*C(I))-
     1      BETA*C(I)**RN*EXP(-BETA*C(I))
            IF(GAMA-1.LT.0.000001) THEN
              CX=C0(I)*ALPHA*(XL(I)/Z(I))/Z(I)
            ELSE
              CX=C0(I)*ALPHA*GAMA*(XL(I)/Z(I))**(GAMA-1)/Z(I)
            ENDIF
            FPRIME(I)=FC*CX
            X(I)=F(I)
 500      CONTINUE
C Find current number of trips
          CALL DISTR(X,OI,DJ,NI,NJ,ERROR,AI,BJ)
C Check if transient stage has passed
          IF(KK.LT.KB) GOTO 5000
C Calculate the Jacobian matrix T
          CALL JACOBN(NI,NJ,OI,DJ,AI,BJ,F,FPRIME,T)
C Multiply matrix T by the previous orthogonal matrix Q
          IFAIL=0
          CALL F01CKF(R,T,Q,LDIM,LDIM,LDIM,ZWK,LDIM,1,IFAIL)
C Do QR factorization of matrix T*Q
          CALL F01QCF(LDIM,LDIM,R,LDIM,ZETA,IFAIL)
          DO 6300 I=1,LDIM
          DO 6300 J=1,LDIM
            Q(I,J)=R(I,J)
 6300     CONTINUE
          CALL F01QEF('S',LDIM,LDIM,LDIM,Q,LDIM,ZETA,WORK,IFAIL)
C
C Calculate Liapunov Exponents and print out
C
```

```
          DO 7000 I=1,LDIM
            CUMR(I)=CUMR(I)+DLOG(ABS(R(I,I)))
 7000     CONTINUE
          IF(MOD(KK,IO).NE.0) GOTO 5000
          DO 8000 I=1,LDIM
            RLE(I)=CUMR(I)/KK
 8000     CONTINUE
          WRITE(*,'(1X,I5,4F18.4)') KK,(X(I),I=1,4)
          WRITE(*,'(1X,I8,4F12.6)') KK,(RLE(I),I=1,LDIM)
          WRITE(1,'(1X,4HRLE=,I8,4F12.6)') KK,(RLE(I),I=1,LDIM)
 5000   CONTINUE
 9000 CONTINUE
      STOP
      END

      SUBROUTINE DISTR (Y,OI,DJ,NI,NJ,E,A,B)
C
C This subroutine solve the doubly constrained gravity
C model by the algorithm described in section 5.3.2.
C
      PARAMETER (NN=30,NZ=5)
      DOUBLE PRECISION Y(NN),OI(NZ),DJ(NZ),E,
     1  X0(NZ,NZ),X(NZ,NZ),A(NZ),B(NZ)
      INTEGER NI,NJ,NIJ,ITER
C Initialize
      DO 200 J=1,NJ
      DO 200 I=1,NI
        X0(I,J)=Y((J-1)*NI+I)
 200  CONTINUE
      DO 250 I=1,NI
        X0(I,NJ+1)=OI(I)
 250  CONTINUE
      DO 300 J=1,NJ+1
        X0(NI+1,J)=DJ(J)
 300  CONTINUE
      ITER=100
      DO 500 J=1,NJ
        B(J)=1.0
 500  CONTINUE
C
C Start the iteration
C
      DO 5000 KK=1,ITER
        LFLAG=(-1)**KK
        IF(LFLAG.EQ.-1) THEN
          DO 600 I=1,NI
            A(I)=0
            DO 550 J=1,NJ
              A(I)=A(I)+B(J)*X0(I,J)
 550        CONTINUE
            A(I)=X0(I,NJ+1)/A(I)
 600      CONTINUE
        ELSE
          DO 700 J=1,NJ
            B(J)=0
            DO 650 I=1,NI
              B(J)=B(J)+A(I)*X0(I,J)
 650        CONTINUE
            B(J)=X0(NI+1,J)/B(J)
 700      CONTINUE
        ENDIF
C Check the convergency
        DO 1000 I=1,NI
        DO 1000 J=1,NJ
          X(I,J)=A(I)*B(J)*X0(I,J)
 1000   CONTINUE
        DO 800 J=1,NJ
        X(NI+1,J)=0
```

```
        DO 800 I=1,NI
          X(NI+1,J)=X(NI+1,J)+X(I,J)
 800    CONTINUE
        DO 900 I=1,NI+1
        X(I,NJ+1)=0
        DO 900 J=1,NJ
          X(I,NJ+1)=X(I,NJ+1)+X(I,J)
 900    CONTINUE
        DO 2000 I=1,NI
          IF(ABS(X(I,NJ+1)-X0(I,NJ+1)).GT.E) GOTO 5000
 2000   CONTINUE
        DO 3000 J=1,NJ
          IF(ABS(X(NI+1,J)-X0(NI+1,J)).GT.E) GOTO 5000
 3000   CONTINUE
        GOTO 5500
 5000 CONTINUE
C Output the result
 5500 DO 6000 J=1,NJ
      DO 6000 I=1,NI
        Y((J-1)*NI+I)=X(I,J)
 6000 CONTINUE
      RETURN
      END

      SUBROUTINE JACOBN(NI,NJ,OI,DJ,AI,BJ,F,DF,T)
C
C This program calculate the Jacobian matrix of the gravity
C model by the algorithm described in section 5.4.3.
C
      PARAMETER (NN=30,NZ=5,LDIM=4,NDAB=54)
      DOUBLE PRECISION OI(NZ),DJ(NZ),AI(NZ),BJ(NZ),F(NN),DF(NN),
     1  DFIJ(NZ,NZ),FIJ(NZ,NZ),AAA(NDAB,NDAB),BBB(NDAB),
     2  DAB(NDAB),JAC(9,9),JAC1(9,3,3),JAC2(4,2,2),T(LDIM,LDIM)
      INTEGER NI,NJ,NIJ,IFAIL
      DOUBLE PRECISION AA(NDAB,NDAB),WKS1(NDAB),WKS2(NDAB),XXX
C
C Initialize the linear equation containing partial derivatives
C of balancing factors with respect to the number of trips
C
      NDIM=NI*NJ
      DO 700 J=1,NJ
      DO 700 I=1,NI
        FIJ(I,J)=F((J-1)*NI+I)
        DFIJ(I,J)=DF((J-1)*NI+I)
 700  CONTINUE
      DO 900 I=1,(NI+NJ)*NDIM
      DO 900 J=1,(NI+NJ)*NDIM
        AAA(I,J)=0.0
 900  CONTINUE
      DO 1000 I=1,NI
      DO 1000 II=1,NDIM
        AAA((I-1)*NDIM+II,(I-1)*NDIM+II)=OI(I)
 1000 CONTINUE
      DO 1100 I=1,NJ
      DO 1100 II=1,NDIM
        AAA(NI*NDIM+(I-1)*NDIM+II,NI*NDIM+(I-1)*NDIM+II)=DJ(I)
 1100 CONTINUE
      DO 1200 I=1,NI
      DO 1200 J=1,NJ
      DO 1200 II=1,NDIM
        AAA((I-1)*NDIM+II,NI*NDIM+(J-1)*NDIM+II)=AI(I)*AI(I)*FIJ(I,J)
 1200 CONTINUE
      DO 1300 J=1,NJ
      DO 1300 I=1,NI
      DO 1300 II=1,NDIM
        AAA(NI*NDIM+(J-1)*NDIM+II,(I-1)*NDIM+II)=BJ(J)*BJ(J)*FIJ(I,J)
 1300 CONTINUE
      DO 1500 I=1,NI
```

```
      DO 1500 L=1,NJ
      DO 1500 K=1,NI
        IF(I.EQ.K) THEN
          BBB((I-1)*NDIM+(L-1)*NJ+K)=-AI(I)*AI(I)*BJ(L)*DFIJ(K,L)
        ELSE
          BBB((I-1)*NDIM+(L-1)*NJ+K)=0
        ENDIF
 1500 CONTINUE
      DO 1600 J=1,NJ
      DO 1600 L=1,NJ
      DO 1600 K=1,NI
        IF(J.EQ.L) THEN
          BBB(NI*NDIM+(J-1)*NDIM+(L-1)*NJ+K)=
     1     -BJ(J)*BJ(J)*AI(K)*DFIJ(K,L)
        ELSE
          BBB(NI*NDIM+(J-1)*NDIM+(L-1)*NJ+K)=0
        ENDIF
 1600 CONTINUE
C
C Solve the linear equation to get partial derivatives
C of balancing factors with respect to the number of trips
C
      IFAIL=0
      CALL F04ATF(AAA,NDAB,BBB,NDAB,DAB,AA,NDAB,WKS1,WKS2,IFAIL)
C
C Calculate elements of Jacobian matrix
C
      DO 3000 J=1,NJ
      DO 3000 I=1,NI
      DO 3000 L=1,NJ
      DO 3000 K=1,NI
        JACI=(J-1)*NI+I
        JACJ=(L-1)*NI+K
        IDA=(I-1)*NDIM+JACJ
        JDB=NI*NDIM+(J-1)*NDIM+JACJ
        IF(K.EQ.I .AND. L.EQ.J) THEN
          JAC(JACI,JACJ)=BJ(J)*FIJ(I,J)*DAB(IDA)+
     1           AI(I)*FIJ(I,J)*DAB(JDB)+
     2           AI(I)*BJ(J)*DFIJ(I,J)
        ELSE
          JAC(JACI,JACJ)=BJ(J)*FIJ(I,J)*DAB(IDA)+
     1           AI(I)*FIJ(I,J)*DAB(JDB)
        ENDIF
 3000 CONTINUE
      DO 3800 II=1,NDIM
        DO 3700 J=1,NJ
        DO 3700 I=1,NI
          JAC1(II,I,J)=JAC(II,(J-1)*NI+I)
 3700   CONTINUE
 3800 CONTINUE
      DO 5050 K=1,NI-1
        DO 5000 J=1,NJ-1
        DO 5000 I=1,NI-1
          JAC2(K,I,J)=JAC1(K,I,J)-JAC1(K,I,3)-JAC1(K,3,J)
 5000   CONTINUE
 5050 CONTINUE
      DO 5150 K=4,5
        DO 5100 J=1,2
        DO 5100 I=1,2
          JAC2(K-1,I,J)=JAC1(K,I,J)-JAC1(K,I,3)-JAC1(K,3,J)
 5100   CONTINUE
 5150 CONTINUE
C
C Output the Jacobian matrix
C
      DO 5500 K=1,LDIM
      DO 5500 J=1,2
      DO 5500 I=1,2
```

```
        T(K,(J-1)*(NI-1)+I)=JAC2(K,I,J)
5500 CONTINUE
     RETURN
     END
```

B.4. THE PROGRAM CORDIM

```
      PROGRAM CORDIM
C
C This program calculates correlation function of an
C attractor by the method due to Grassberger and
C Procaccia (1983).  The method is described in
C section 5.5.1.  The logarithms of both correlation
C function and corresponding distance are produced;
C which are ready to be used to obtain the correlation
C dimension of the attractor by, for example, MATLAB
C (The MathWorks Inc. 1993).  The attractor is defined
C by a time series data.
C
C Input data:
C NPT   = length of time series data;
C NDIM  = dimension of the phase space;
C Time series data is in the data file TSDATA.DAT.
C
C Output data:
C LOGR(K) = logarithm of the Kth distance;
C LOGC(K) = logarithm of the Kth value of correlation function.
C
      PARAMETER (LDIM=10,LDAT=10000,LSTEP=200)
      INTEGER NDIM,NPT,NR(LSTEP),NSTEP,OFFSET,MM,EMIN,EMAX,
     1  KMAX,KMIN,KDEL,INDX
      DOUBLE PRECISION X(LDAT,LDIM),LOGC(LSTEP),LOGR(LSTEP),
     1  BASE,DIST,EXPNT,CR
      OPEN(1,FILE='TSDATA',STATUS='OLD',FORM='FORMATTED')
      OPEN(2,FILE='DIMDATA',STATUS='UNKNOWN',FORM='FORMATTED')
C
C Initialization
C
      BASE=2.0
      MM=1
      EMIN=(-50)*MM
      EMAX=20*MM
      NSTEP=EMAX-EMIN+1
      OFFSET=-EMIN+1
      KMIN=NSTEP
      KMAX=1
      KDEL=0
      NDIM=2
      NPT=20000
C
C Read in the time series data
C
      DO 50 I=1,NPT
        READ (1,*) KKK,(X(I,J),J=1,NDIM)
 50   CONTINUE
C
C Set the vector for counting numbers of inter-point
C
      DO 500 I=1,NSTEP
        NR(I)=0.0
 500  CONTINUE
C
C Calculate the distance of each inter-point
C
      DO 3000 II=1,NPT-1
      DO 3000 JJ=II+1,NPT
        DIST=0
        DO 1000 I=1,NDIM
```

```
          DIST=DIST+(X(JJ,I)-X(II,I))*(X(JJ,I)-X(II,I))
 1000   CONTINUE
        DIST=SQRT(DIST)
        IF(DIST.EQ.0) THEN
          KDEL=KDEL+1
          GOTO 3000
        ENDIF
        EXPNT=FLOAT(MM)*DLOG(DIST)/DLOG(BASE)
        INDX=INT(EXPNT)+OFFSET
        NR(INDX)=NR(INDX)+1
        IF(INDX.LT.KMAX) KMAX=INDX
        IF(INDX.GT.KMIN) KMIN=INDX
 3000 CONTINUE
C
C Calculate the number of points within the given distances
C
      DO 4000 K=2,NSTEP
        NR(K)=NR(K-1)+NR(K)
 4000 CONTINUE
C
C Find the logarithms of both correlation function and
C the distance
C
      DO 4100 K=1,NSTEP
        IF(NR(K).EQ.0) THEN
          LOGC(K)=0.0
        ELSE
          CR=FLOAT(2*NR(K))/FLOAT((NPT*NPT))
          LOGC(K)=DLOG(CR)/DLOG(BASE)
        ENDIF
        LOGR(K)=FLOAT((K-OFFSET))/FLOAT(MM)
        WRITE(*,'(1X,I4,2F20.4)') K,LOGR(K),LOGC(K)
        WRITE(2,'(1X,I4,2F20.4)') K,LOGR(K),LOGC(K)
 4100 CONTINUE
      STOP
      END
```

B.5. THE PROGRAM TSLE

```
      PROGRAM TSLE
C
C This program calculates Liapunov exponents of an attractor
C reconstructed from time series data of a single variable.
C The algorithm used is due to Eckmann et al. (1986).
C The program uses NAG (NAG Ltd, 1987) routines:
C M01DAF, M01ZAF, F04JGF, F01CKF, F01QCF, and F01QEF.
C
C Input data:
C LTS    = length of time series data;
C EDIM   = embedding dimension;
C TAU    = delay time used to reconstruct the attractor;
C IO     = time interval for print out.
C The time series data is in the data file TSDATA.DAT.
C
C Output data:
C ELE(I) = the Ith Liapunov exponent of the
C reconstructed attractor.
C
      PARAMETER (LDIM1=75,LDIM2=10,LDIM3=10000)
      INTEGER RPERM(LDIM3),IPERM(LDIM3),IPLS(LDIM1),EDIM,TAU
      DOUBLE PRECISION OTS(LDIM3),OBSV(LDIM1,LDIM2),OBSV1(LDIM1),
     1  T(LDIM2,LDIM2),Q(LDIM2,LDIM2),R(LDIM2,LDIM2),
     2  CUMR(LDIM2),ELE(LDIM2),DT,MYDAT(4),
     3  WORK(75),Z(LDIM2),ZETA(LDIM2),SIGMA
      LOGICAL SVD
      RCA(I,J)=OTS(I+(J-1)*TAU)
      OPEN(1,FILE='TSDATA',STATUS='OLD',FORM='FORMATTED')
      OPEN(2,FILE='LEDATA',STATUS='UNKNOWN',FORM='FORMATTED')
C
C Input data
C
      PRINT*,'Input the length of time series'
      READ*, LTS
      PRINT*,'Input the embedding dimension'
      READ*, EDIM
      PRINT*,'Input the delay for reconstruction'
      READ*, TAU
      PRINT*,'Input the interval for output'
      READ*, IO
      DO 10 I=1,LTS
        READ (1,*) OTS(I)
 10   CONTINUE
C
C Initialize
C
      LSTEP=TAU
      BALL0=0.001
      DBALL=0.0001
      BALLM=1.0
      LSMAX0=3*EDIM
      LSMAXD=2
      LSMAX1=6*EDIM
      LSMIN=2*EDIM
      DT=1
C
C Sort the time series data in the increasing order
C
      IFAIL=0
      CALL M01DAF(OTS,1,LTS,'A',RPERM,IFAIL)
      DO 30 I=1,LTS
```

```
        IPERM(I)=RPERM(I)
 30   CONTINUE
      CALL M01ZAF(IPERM,1,LTS,IFAIL)
C
C Calculate number of points on the reconstructed attractor
C
      NPT=LTS-(EDIM-1)*TAU-LSTEP
C
C Initialize for Jacobian matrix estimation
C
      DO 80 I=1,EDIM
        DO 50 J=1,EDIM
          T(I,J)=0.0
          Q(I,J)=0.0
 50     CONTINUE
        CUMR(I)=0.0
        Q(I,I)=1.0
 80   CONTINUE
      DO 90 I=1,EDIM-1
        T(I,I+1)=1.0
 90   CONTINUE
C
C Begin the journy through the reconstructed attractor
C
      DO 10000 IIII=1,NPT,LSTEP
        NPLS=0
        LSMAX=LSMAX0
        BALL=BALL0
        IRNK=RPERM(IIII)
C
C Find the neighbours of the two current points
C
        IRNKF=IRNK+1
        IRNKB=IRNK-1
 500    DO 1000 III=IRNKF,LTS
          IF(NPLS.GE.LSMAX) GOTO 5000
          ICN=IPERM(III)
          IF(ICN.GT.NPT) GOTO 1000
          DO 600 II=1,NPLS
            IF(ICN.EQ.IPLS(II)) GOTO 1000
            ICN1=IPLS(II)
            DO 550 I=1,EDIM
              DIST=ABS(RCA(ICN,I)-RCA(ICN1,I))
              IF(DIST.GT.0.0000001) GOTO 600
 550        CONTINUE
            GOTO 1000
 600      CONTINUE
          DO 800 II=1,EDIM
            DIST=ABS(RCA(ICN,II)-RCA(IIII,II))
            IF(DIST.LE.BALL) THEN
              GOTO 800
            ELSE IF (II.EQ.1)THEN
              GOTO 1500
            ELSE
              GOTO 1000
            ENDIF
 800      CONTINUE
          DIST=ABS(RCA(ICN+LSTEP,EDIM)-RCA(IIII+LSTEP,EDIM))
          IF(DIST.GT.BALL) GOTO 1000
          NPLS=NPLS+1
          IPLS(NPLS)=ICN
 1000   CONTINUE
 1500   DO 2000 III=IRNKB,1,-1
          IF(NPLS.GE.LSMAX) GOTO 5000
          ICN=IPERM(III)
          IF(ICN.GT.NPT) GOTO 2000
          DO 1600 II=1,NPLS
            IF(ICN.EQ.IPLS(II)) GOTO 2000
```

```
            ICN1=IPLS(II)
            DO 1550 I=1,EDIM
              DIST=ABS(RCA(ICN,I)-RCA(ICN1,I))
              IF(DIST.GT.0.0000001) GOTO 1600
 1550       CONTINUE
            GOTO 2000
 1600     CONTINUE
          DO 1800 II=1,EDIM
            DIST=ABS(RCA(ICN,II)-RCA(IIII,II))
            IF(DIST.LE.BALL) THEN
              GOTO 1800
            ELSE IF (II.EQ.1)THEN
              GOTO 3000
            ELSE
              GOTO 2000
            ENDIF
 1800     CONTINUE
          DIST=ABS(RCA(ICN+LSTEP,EDIM)-RCA(IIII+LSTEP,EDIM))
          IF(DIST.GT.BALL) GOTO 2000
          NPLS=NPLS+1
          IPLS(NPLS)=ICN
 2000   CONTINUE
C
C Check the number of points for regression
C
 3000   IF(NPLS.GE.LSMAX) GOTO 5000
C Increase the size of ball to get more neighbours
        BALL=BALL+DBALL
        IF(BALL.LT.BALLM) GOTO 500
        IF(NPLS.LT.LSMIN .OR. LSMAX.GT.LSMAX0) THEN
          PRINT*,'NO ENOUGH NEIBOURS!BALL,LSMAX,NPLS=',
     *                          BALL,LSMAX,NPLS
          STOP
        ENDIF
        GOTO 5000
C
C Estimate the Jacobian matrix T by Least Squares
C
 5000   DO 5500 I=1,NPLS
          IND=IPLS(I)
          DO 5300 J=1,EDIM
            OBSV(I,J)=RCA(IND,J)-RCA(IIII,J)
 5300     CONTINUE
          OBSV1(I)=RCA(IND+LSTEP,EDIM)-RCA(IIII+LSTEP,EDIM)
 5500   CONTINUE
        CALL F04JGF(NPLS,EDIM,OBSV,LDIM1,OBSV1,5*0.00001,SVD,SIGMA,
     *          IRANK,WORK,75,IFAIL)
C Check the rank of T
        DO 5550 I=1,EDIM
          IF(ABS(OBSV1(I)).LT.0.00000000000001) THEN
          LSMAX=LSMAX+LSMAXD
          IF(LSMAX.LE.LSMAX1) GOTO 500
            PRINT*,'T is singular!BALL,LSMAX,NPLS=',BALL,LSMAX,NPLS
            STOP
          ENDIF
 5550   CONTINUE
C
C Construct the Jacobian matrix T
C
        DO 5800 I=1,EDIM
          T(EDIM,I)=OBSV1(I)
 5800   CONTINUE
C
C Multiply matrix T by the previous orthogonal matrix Q
C
        CALL F01CKF(R,T,Q,LDIM2,LDIM2,LDIM2,Z,LDIM2,1,IFAIL)
C
C Do QR factorization of the matrix T'= T*Q
```

```
C
        CALL F01QCF(EDIM,EDIM,R,LDIM2,ZETA,IFAIL)
        DO 6300 I=1,EDIM
        DO 6300 J=1,EDIM
          Q(I,J)=R(I,J)
 6300   CONTINUE
        DO 6400 I=2,EDIM
        DO 6400 J=1,I-1
          R(I,J)=0.0
 6400   CONTINUE
        CALL F01QEF('S',EDIM,EDIM,EDIM,Q,LDIM2,ZETA,WORK,IFAIL)
C
C Calculate Liapunov Exponents and print out
C
        DO 7000 I=1,EDIM
          CUMR(I)=CUMR(I)+DLOG(ABS(R(I,I)))
 7000   CONTINUE
        IF(MOD(IIII/LSTEP,IO).NE.0) GOTO 10000
        DO 8000 I=1,EDIM
          ELE(I)=CUMR(I)/(IIII*DT)
 8000   CONTINUE
        WRITE(*,'(1X,I7,9F8.3)') IIII,(ELE(I),I=1,EDIM)
        WRITE(2,'(1X,I7,9F8.3)') IIII,(ELE(I),I=1,EDIM)
10000 CONTINUE
      STOP
      END
```

B.6. THE PROGRAM TSDIM

```
      PROGRAM TSDIM
C
C This program calculates the correlation function of an
C attractor by the method due to Grassberger and
C Procaccia (1983).  The logarithms of both correlation
C function and the corresponding distance are produced;
C which are ready to be used to obtain the correlation
C dimension of the attractor by, for example, MATLAB
C (The MathWorks Inc. 1993).  The attractor is
C reconstructed from time series data of a single variable
C of a dynamical system.
C
C Input data:
C LTS   = length of time series data;
C NDIM  = embedding dimension;
C TAU   = delay time used to reconstruct the attractor;
C The time series data is in the data file TSDATA.DAT.
C
C Output data:
C LOGR(K) = logarithm of the Kth distance;
C LOGC(K) = logarithm of the Kth value of correlation function.
C
      PARAMETER (LDIM=10,LDAT=10000,LSTEP=200)
      INTEGER LTS,TAU,NDIM,NPT,NR(LSTEP),NSTEP,OFFSET,MM,
     1  EMIN,EMAX,KMAX,KMIN,KDEL,INDX
      DOUBLE PRECISION OTS(LDAT),X(LDAT,LDIM),LOGC(LSTEP),
     1  LOGR(LSTEP),BASE,DIST,EXPNT,CR,MYDAT(4)
      OPEN(1,FILE='TSDATA',STATUS='OLD',FORM='FORMATTED')
      OPEN(2,FILE='DIMDATA',STATUS='UNKNOWN',FORM='FORMATTED')
C
C Initialize
C
      BASE=2.0
      MM=1
      EMIN=(-50)*MM
      EMAX=20*MM
      NSTEP=EMAX-EMIN+1
      OFFSET=-EMIN+1
      KMIN=NSTEP
      KMAX=1
      KDEL=0
      PRINT*,'Input the length of time series'
      READ*, LTS
      PRINT*,'Input the embedding dimension'
      READ*, NDIM
      PRINT*,'Input the delay for reconstruction'
      READ*, TAU
C
C Read in the time series data
C
      DO 10 I=1,LTS
        READ (1,*) OTS(I)
 10   CONTINUE
C
C Calculate the number of points on the reconstructed attractor
C
      NPT=LTS-(NDIM-1)*TAU
      DO 50 I=1,NPT
      DO 50 J=1,NDIM
        X(I,J)=OTS(I+(J-1)*TAU)
 50   CONTINUE
```

```
C
C Set the vector for counting the numbers of inter-point
C
      DO 500 I=1,NSTEP
        NR(I)=0
 500  CONTINUE
C
C Calculate the distances of every inter-point
C
      DO 3000 II=1,NPT-1
      DO 3000 JJ=II+1,NPT
        DIST=0
        DO 1000 I=1,NDIM
          DIST=DIST+(X(JJ,I)-X(II,I))*(X(JJ,I)-X(II,I))
 1000   CONTINUE
        DIST=SQRT(DIST)
        IF(DIST.EQ.0) THEN
          KDEL=KDEL+1
          GOTO 3000
        ENDIF
        EXPNT=FLOAT(MM)*DLOG(DIST)/DLOG(BASE)
        INDX=INT(EXPNT)+OFFSET
        NR(INDX)=NR(INDX)+1
        IF(INDX.LT.KMAX) KMAX=INDX
        IF(INDX.GT.KMIN) KMIN=INDX
 3000 CONTINUE
C
C Calculate the number of points within the given distance
C
      DO 4000 K=2,NSTEP
        NR(K)=NR(K-1)+NR(K)
 4000 CONTINUE
C
C Find the logrithms of both the correlation function
C and the distance
C
      DO 4100 K=1,NSTEP
        IF(NR(K).EQ.0) THEN
          LOGC(K)=0.0
        ELSE
          CR=FLOAT(2*NR(K))/FLOAT((NPT*NPT))
          LOGC(K)=DLOG(CR)/DLOG(BASE)
        ENDIF
        LOGR(K)=FLOAT((K-OFFSET))/FLOAT(MM)
        WRITE(*,'(1X,I4,2F20.4)') K,LOGR(K),LOGC(K)
        WRITE(2,'(1X,I4,2F20.4)') K,LOGR(K),LOGC(K)
 4100 CONTINUE
      STOP
      END
```

REFERENCES

BAKER RGV (1983). "On the kinematics and quantum dynamics of traffic flow", *Transportation Research*, **17B**(1), 55–66

BARTLETT MS (1990). "Chance or chaos?", *Journal of the Royal Statistical Society*, 153**A**(3), 321–347

BECKMANN M, MCGUIRE CB, & WINSTEN C (1956). *Studies in the Economics of Transportation*, Yale University Press, New Haven, Conn

BENDER JG & FENTON RE (1970). "On the flow capacity of automated highways", *Transportation Science*, **4**(1), 52–63

BRANSTON D (1976). "Link capacity functions: a review", *Transportation Research*, **10**(4), 223–236

CHANDLER RE, HERMAN R, & MONTROLL EW (1958). "Traffic dynamics: studies in car following", *Operations Research*, **6**, 165–184

CONTE R & DUBOIS M (1988). "Liapunov exponents of experimental systems", *Nonlinear Evolutions* (LEON JJP, Editor), World Scientific, 767–779

CONTE SD & BOOR CD (1980). *Elementary Numerical Analysis, an Algorithmetic Approach*, McGraw-Hill, Inc.

DENDRINOS DS & SONIS M (1990). *Chaos and Social-Spatial Dynamics*, Springer-Verlag

DIAL RB (1971). "A probabilistic multipath traffic assignment model which obviates path enumeration", *Transportation Research*, **5**(2), 83–111

DISBRO JE & FRAME M (1990). "Traffic flow theory and chaotic behaviour", *Transportation Research Record*, **1225**, 109–115

DRIVER RD (1977). *Ordinary and Delay-differential Equations*, Springer-Verlag

ECKMANN J-P, KAMPHORST SO, RUELLE D, & SILIBERTO (1986). "Liapunov exponents from time series", *Physical Review* **34A**, 4971–4979

ECKMANN J-P & RUELLE D (1985). "Ergodic theory of chaos and strange attractors", *Reviews of Modern Physics* **57**(3), 617–656

ERLANDER S (1990). "Efficient population behaviour and the simultaneous choices of origins, destinations and routes", *Transportation Research*, 24B(5), 363–373

ERLANDER S & STEWART NF (1990). *The Gravity Model in Transportation Analysis — Theory and Extensions*, VSP, Utrecht, The Netherlands

EVANS SP (1973). "A relationship between the gravity model for trip distribution and the transportation problem in linear programming", *Transportation*

Research, **7**, 39–61

EVANS SP (1976). "Derivation and analysis of some models for combined trip distribution and assignment", *Transportation Research*, **10**, 37–57

FARMER JD (1982). "Chaotic attractors of an infinite-dimensional dynamical system", *Physica*, **7D**, 366–393

FERNANDEZ JE & FRIESZ TL (1983). "Equilibrium predictions in transportation markets: the state of the art", *Transportation Research*, **17B**(2), 155–172

FERRARI P (1994). "The instability of motorway traffic", *Transportation Research*, **28B**(2), 175–186

FISK C (1980). "Some developments in equilibrium traffic assignment", *Transportation Research*, **14B**, 243–255

FLORIAN M, & NGUYEN S (1978). "A combined trip distribution modal split and trip assignment model ", *Transportation Research*, **12**, 241–246

FRANK M & STENGOS T (1988). "Chaotic dynamics in economic time-series", *Journal of Economic Surveys*, **2**(2), 103–113

FRIESZ TL (1985). "Transportation network equilibrium, design and aggregation: key developments and research opportunities", *Transportation Research*, **19A**(5/6), 413–427

GAZIS DC, HERMAN R, & ROTHERY RW (1961). "Nonlinear follow-the-leader models of traffic flow", *Operations Research*, **9**(4), 545–567

GLEICK J (1987). *Chaos: Making a New Science*, New York: Viking

GRASSBERGER P & PROCACCIA I (1983). "Measuring the strangeness of strange attractors", *Physica* 9*D*, **9*D***, 189–208

GRIFFEL DL (1981). *Applied Functional Analysis*, Ellis Horwood Ltd

GYORI I & LADAS G (1991). *Oscillation Theory of Delay Differential Equations: with Applications*, Clarendon press, Oxford

HAGER WW (1988). *Applied Numerical Linear Algebra*, Prentice-Hall International, Inc.

HALE JK & LUNEL (1993). *Introduction to Functional Differential Equations*, Springer-Verlag

HENON M (1976). "A two-dimensional mapping with a strange attractor", *Communication of Mathematical Physics*, **50**

HERMAN R, MONTROLL EW, POTTS RB & ROTHERY RW (1959). "Traffic dynamics: analysis of stability in car following", *Operations Research*, **7**, 86–106

HERMAN R & POTTS RB (1961). "Single-lane traffic theory and experiment", *Theory of traffic flow, Proceedings of the Symposium on the Theory of traffic flow* (HERMAN R, Editor), Warren.1961.

HOROWITZ JL (1984). "The stability of stochastic equilibrium in a two-link transportation network", *Transportation Research*, 18B, 13–28

JARRETT DF & ZHANG XY (1993). "The dynamic behaviour of road traffic flow: stability or Chaos?", *BCS Displays Group: Application of Fractals and Chaos* (CRILLY AJ, EARNSHAW RA, & JONES H, Editors:), Springer-Verlag, 237–248

KAPUR KC (1971). "Vehicle-following behaviour by calculus of variations", Transportation Research, **5**, 69–73

KIRBY HR & SMITH MJ (1991). "Can chaos theories have transport applications", *23rd Annual conference of the Universities Transport Study Group*, University of Nottingham (unpublished)

KUANG Y (1993). *Delay Differential Equations with Applications in Population Dynamics*, Academic press, Inc.

KUHNE RD (1987). "Freeway speed distribution and acceleration noise-calculations from a stochastic continuum theory and comparison with measurements", Transportation and Traffic Theory, *Proceedings of the 10th International Symposium on Transportation and Traffic Theory* (GARTNER NH & WILSON NHM, Editors), Elsevier, New York, 1987, 119–137

KUHNE RD (1991). "Chaotic behaviour of traffic flow", Preprint

KUHNE RD & BECKSCHULTE R (1993). "Non-linearity stochastics of unstable traffic flow", *Transportation and Traffic Theory, Proceedings of the 12th International Symposium on the Theory of Traffic flow and Transportation* (DAGANZO CF, Editor), Elsevier, New York, 367–386

LAM WHK, & HUANG HJ (1992). "A combined trip distribution and assignment model for multiple user classes", *Transportation Research*, **26B**(4), 275–287

LANCASTER P (1969). *Theory of Matrices*, Academic press, New York, London

LEUTZBACH W (1988). *Introduction to the Theory of Traffic Flow*, Springer-Verlag

LIGHTHILL MJ & WHITHAM FB (1955). "On kinematic waves: II. A theory of traffic flow on long, crowded roads", *Proceedings of Royal Society* (series A), **229**, 317–345

LORENZ E N (1963). "Deterministic nonperiodic flow", *Journal of Atmosphere Science* **20**

NAG (1987). *NAG Fortran Library Manual*, Mark 12, NAG Limited

NEWELL GF (1982). *Applications of queueing theory*, London; Chapman and Hall

ORTUZAR JdeD & WILLUMSEN LG (1990). *Modelling Transport*, Wiley

PACKARD NH, CRUTCHFIELD JP, FARMER JD, & SHAW RS (1980). "Geometry from a time series", *Physical Review Letters*, **45**(9), 712–716

PARKER TS & CHUA LO (1989). *Practical Numerical Algorithms for Chaotic Systems,* Springer-Verlag

PAYNE HJ (1979). "A critical review of a macroscopic freeway model", *Research Directions in Computer Control of Urban traffic Systems* (ASCE), New York, 251–265

PETERS J (1993). *Undergraduate Dissertation*, BA Accounting and Finance, Middlesex University

POTTS RB & OLIVER RM (1972). *Flows in Transportation networks*, Academic press

ROSS P (1988). "Some properties of macroscopic traffic models", *Transportation Research Record*, **1194**, 129–134

ROUX JC, SIMOYI RH, & SWINNEY HL (1983). "Observation of a strange attractor", *Physica* **8D**, 257–266

SHEFFI Y (1985). *Urban Transportation Networks: Equilibrium Analysis with Mathematical Programming Methods*, Prentice Hall, Englewood, Cliffs, New Jersey

SMITH MJ (1984). "The stability of a dynamic model of traffic assignment—an application of a method of Liapunov", *Transportation Science*, **18**(3), 245–252

SMITH MJ & GHALI MO (1991). "A new dynamic model for evaluating the performance for urban traffic control systems and route guidance strategies", *Transportation Research Record*, **1036**, 33–39

STEWART I (1989). *Does God Play Dice? The Mathematics of Chaos*, Oxford: Basil Blackwell

SWINNEY HL (1983). "Observations of order and chaos in nonlinear systems", *Physica* **7D**, 3–15

TCI Software Research (1984–1989). T^3 *Manual*, T^3 Scientific Word Processing System, TCI Software Research

The Department of Transport and the Central Office of Information (1988). *The Highway Code*. HMSO

The MathWorks Inc. (1993). *MATLAB User's Guide.*

THOMPSON JMT & STEWART HB (1988). *Nonlinear Dynamics and Chaos: Geometrical Methods for Engineers and Scientists*, Wiley

UNLU V (1995). *Undergraduate Dissertation*, BA European Business Administration, Middlesex University

UNWIN EA & DUCKSTEIN L (1967). "Stability of reciprocal-spacing type car-following models", *Transportation Science*, **1**, 95–109

VAN VLIET D (1982). "SATURN: a modern assignment model", *Traffic Engineering and Control*, 23(12), 578–581

WARDROP JG (1952). "Some theoretical aspects of road traffic research", *Proceedings of the Institution of Civil Engineers*, Part II, 1(2), 325–378

WATLING D & VAN VUREN T (1993). "The modelling of dynamic route guidance systems", *Transportation Research*, **1C**(2), 159–182

WILHELM WE & SCHMIDT JW (1973). "Review of car following theory",

Transportation Engineering Journal (ASCE), **99**, TE4, 923–933

WOLF A, SWIFT JB, SWINNEY HL, & VASTANO JA (1985). "Determining Liapunov exponents from a time series", *Physica*, **16D**, 285–317

ZHANG XY (1994). "Traffic dynamics: studies in the gravity model and the trip assignment model", *26th Annual conference of the Universities Transport Study Group*, University of Leeds (unpublished)

ADDITIONAL GRAPHS RELATING TO CHAPTERS 4 AND 5

During this research, many more graphs were produced in the numerical analysis of the three traffic models in an attempt to identify all possible types of dynamic behaviour in the models. Only selected ones were shown in the thesis because of the limited space. Here are some additional interesting diagrams, including three periodic solutions of the car-following model, three bifurcation diagrams of the unconstrained or singly constrained gravity model, and four bifurcation diagrams of the doubly constrained gravity model. These diagrams were produced in the same settings as those in Chapters 4 and 5. The notations are the same as well, except that here x1 is the number of trips from origin one to destination one for Figures 4–10. The values of uncongested travel costs and capacities for Figures 4–6 are

$$\mathbf{c}^0 = (c_1^0\ c_2^0\ c_3^0\ c_4^0) = (1.4\ 1.2\ 1.8\ 1.6) \quad \text{and}$$

$$\mathbf{q} = (q_1\ q_2\ q_3\ q_4) = (0.17\ 0.15\ 0.25\ 0.23)$$

respectively. While for Figures 7–11,

$$\mathbf{c}^0 = \begin{bmatrix} c_{11}^0 & c_{12}^0 & c_{13}^0 \\ c_{21}^0 & c_{22}^0 & c_{23}^0 \\ c_{31}^0 & c_{32}^0 & c_{33}^0 \end{bmatrix} = \begin{bmatrix} 1.00 & 1.00 & 1.50 \\ 1.20 & 1.40 & 1.80 \\ 1.50 & 0.90 & 0.50 \end{bmatrix},$$

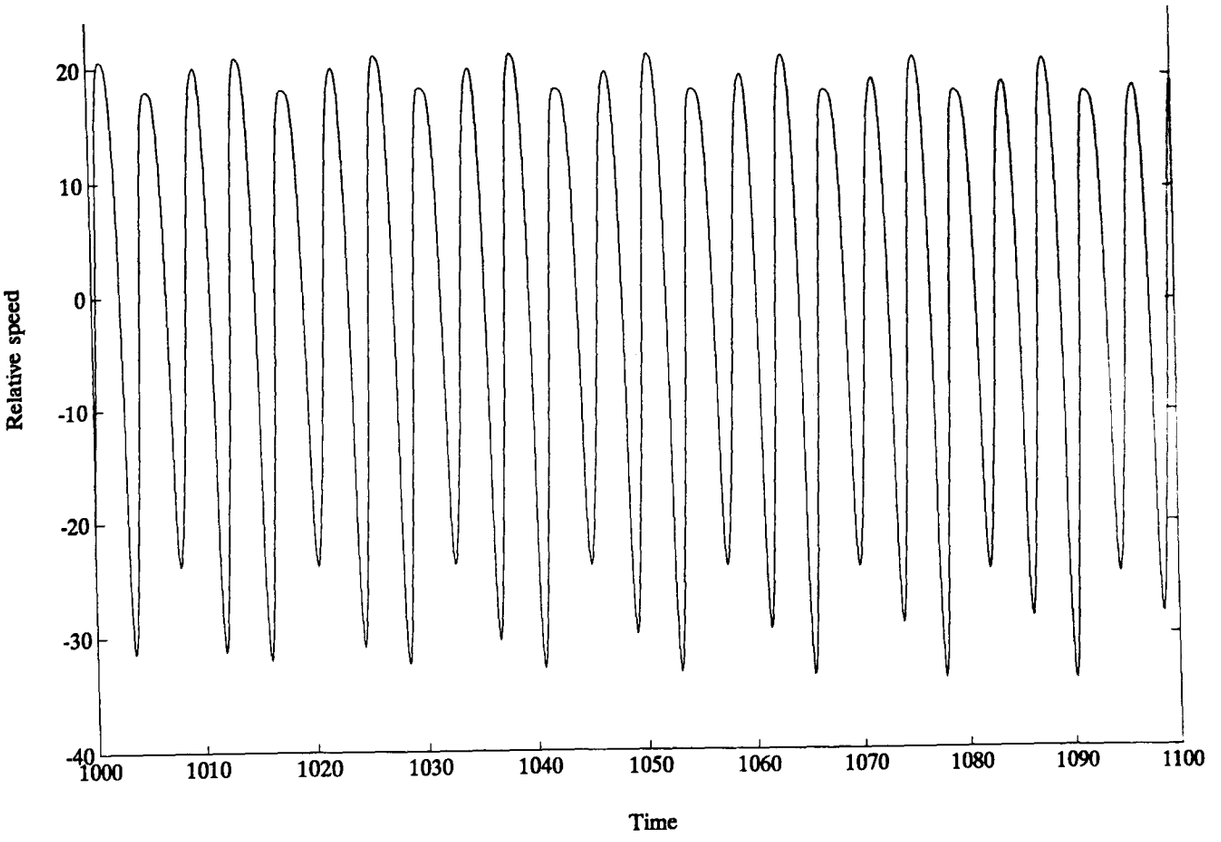
and

$$\mathbf{q} = \begin{bmatrix} q_{11} & q_{12} & q_{13} \\ q_{21} & q_{22} & q_{23} \\ q_{31} & q_{32} & q_{33} \end{bmatrix} = \begin{bmatrix} 0.01 & 0.11 & 0.07 \\ 0.09 & 0.09 & 0.11 \\ 0.10 & 0.09 & 0.07 \end{bmatrix}.$$

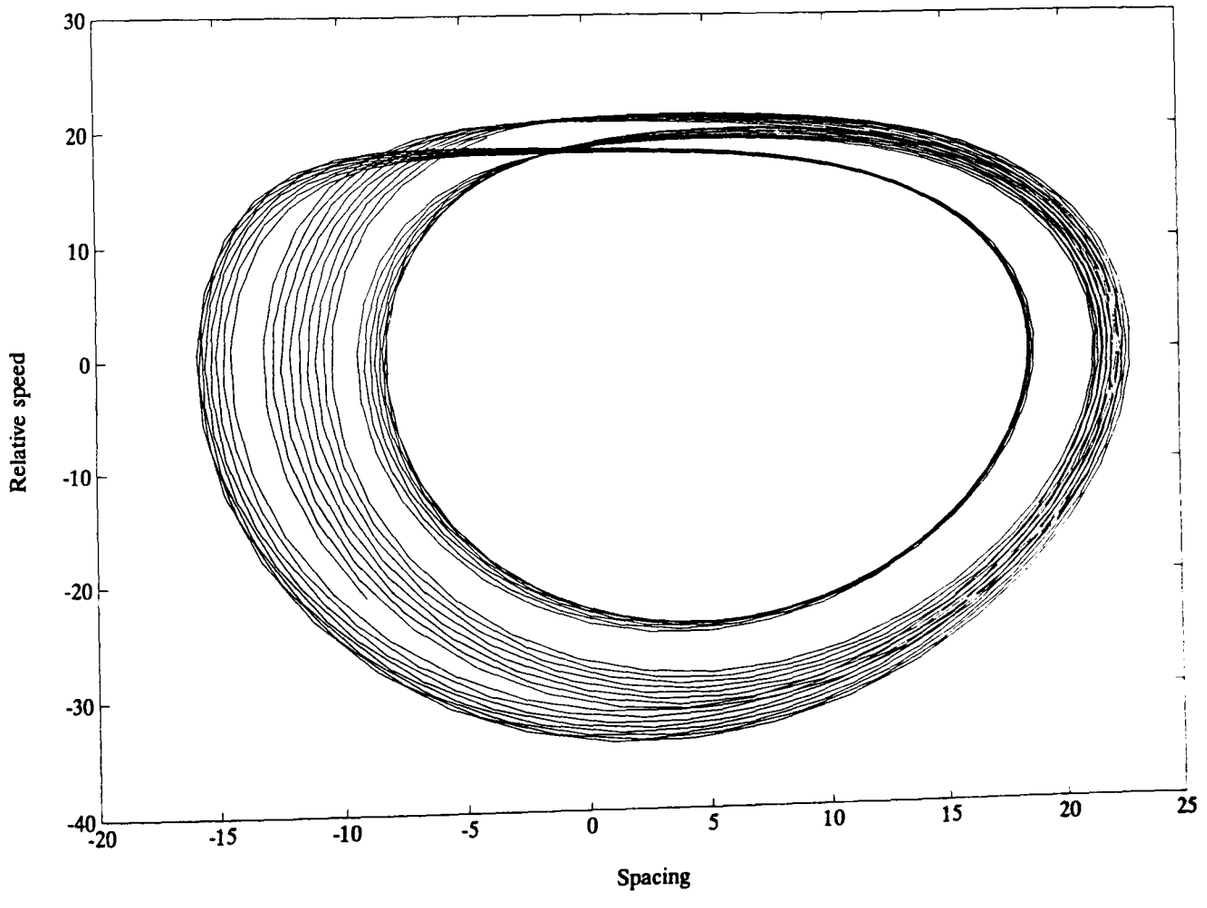
The total numbers of trips from and to each zone are

$\mathbf{o} = (o_1\ o_2\ o_3) = (0.35\ 0.35\ 0.30)$, and $\mathbf{d} = (d_1\ d_2\ d_3) = (0.30\ 0.30\ 0.40)$,

respectively for Figures 7–10.

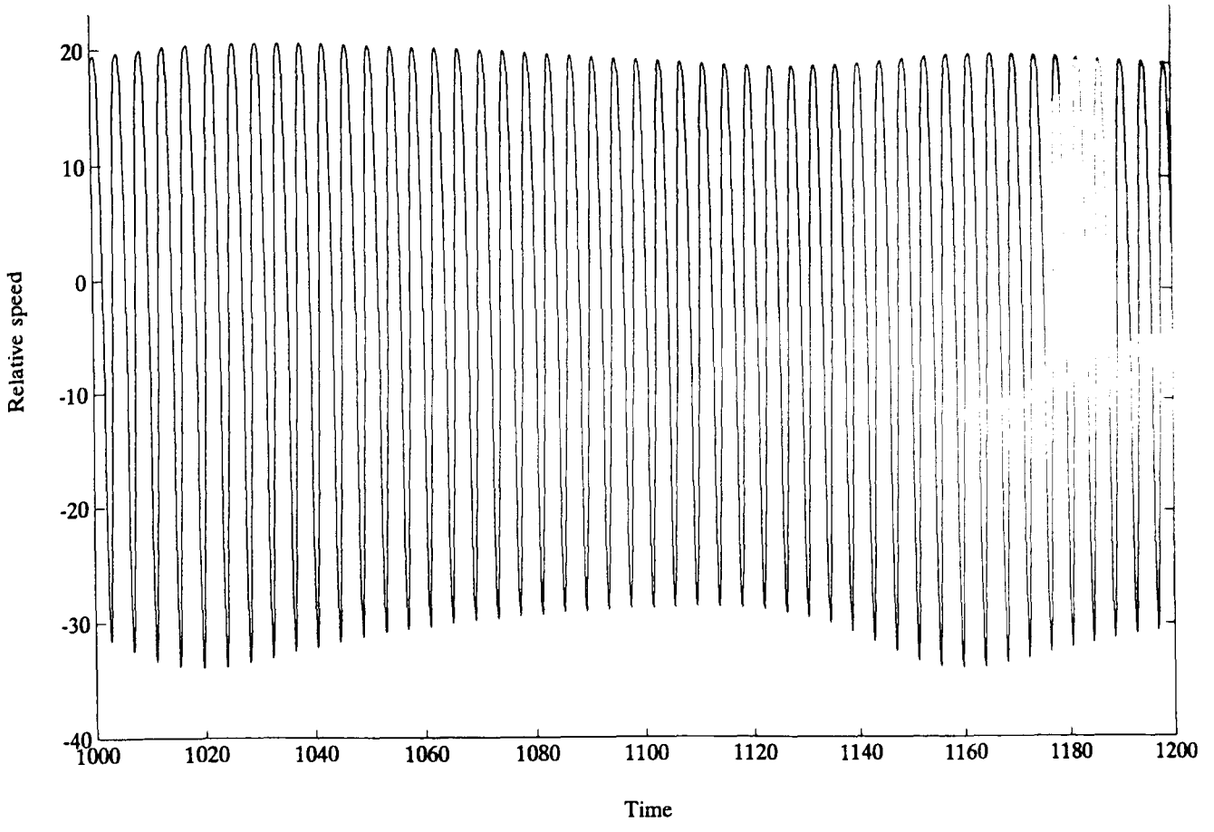


(a)

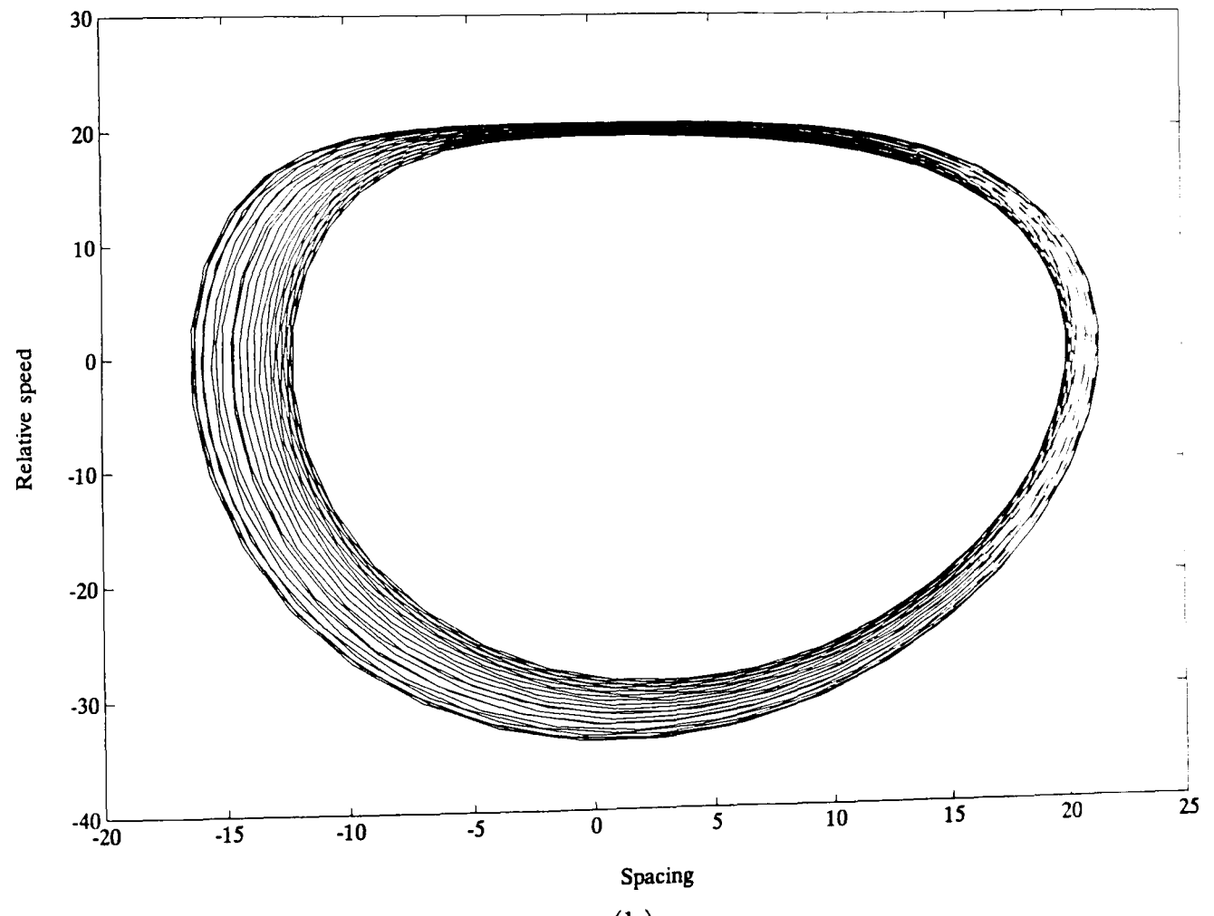


(b)

Figure 1. Periodic solution of the nonlinear non-autonomous car-following model, with $N = 2$, $l = 2.0$, $m = 1.0$, $\alpha = 140$, $\tau = 1.0$, $b = 30$, $F = 1.0$, $\rho = 0.5$, $\bar{u} = 20$. (a) Relative speed; (b) relative speed against spacing.

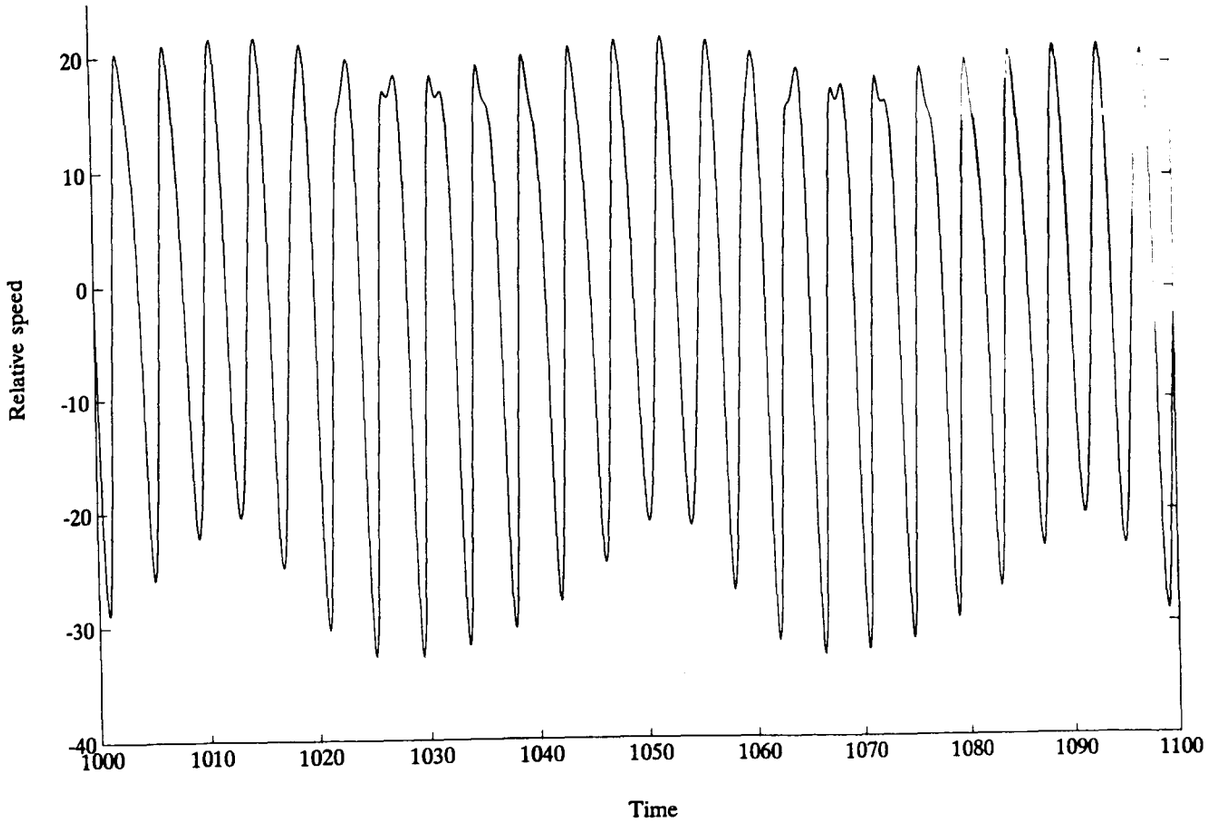


(a)

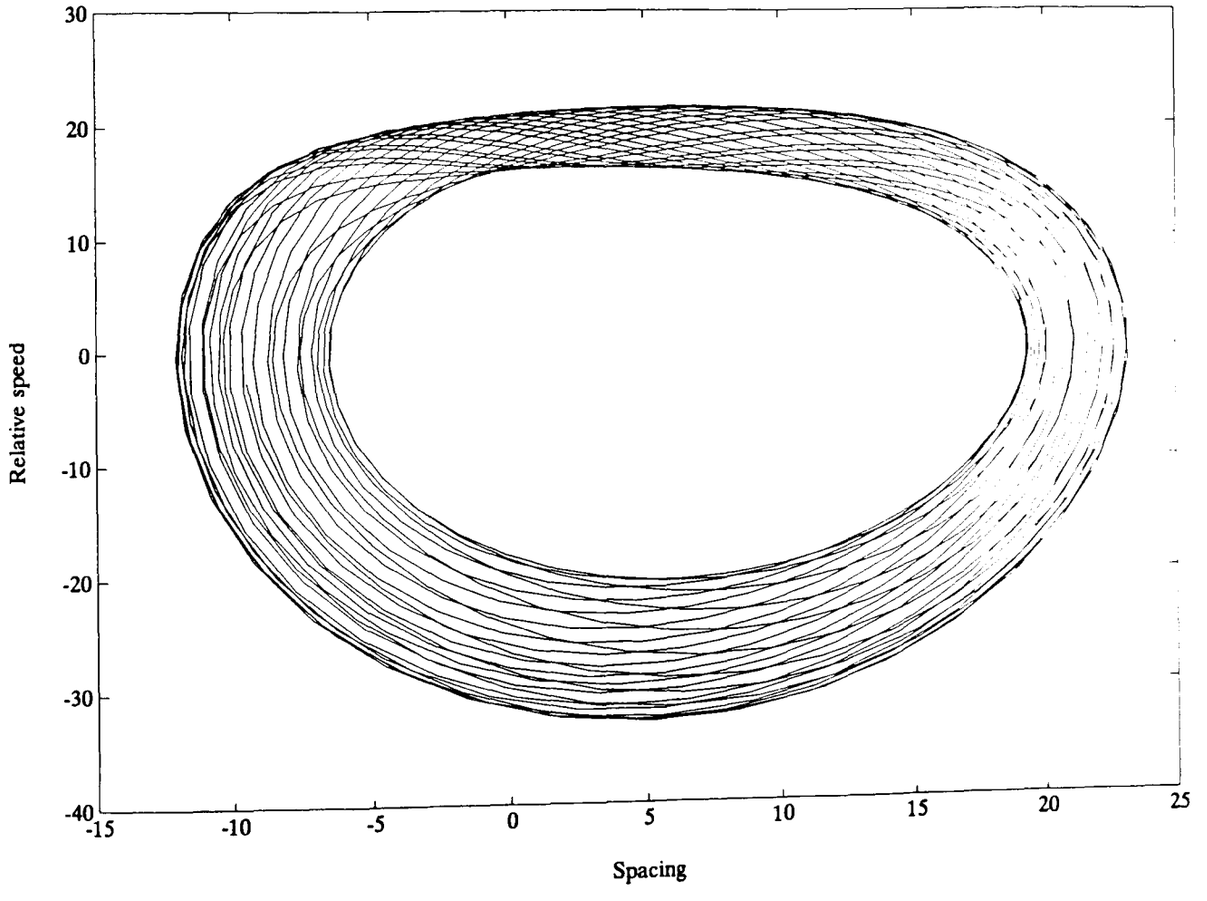


(b)

Figure 2. Periodic solution of the nonlinear non-autonomous car-following model, with $N = 2$, $l = 2.0$, $m = 1.0$, $\alpha = 140$, $\tau = 1.0$, $b = 30$, $F = 1.0$, $\rho = \pi/2$, $\bar{u} = 20$. (a) Relative speed; (b) relative speed against spacing.



(a)



(b)

Figure 3. Periodic solution of the nonlinear non-autonomous car-following model, with $N = 2$, $l = 2.0$, $m = 1.0$, $\alpha = 140$, $\tau = 1.0$, $b = 30$, $F = 10.0$, $\rho = 2.9$, $\bar{u} = 20$. (a) Relative speed; (b) relative speed against spacing.

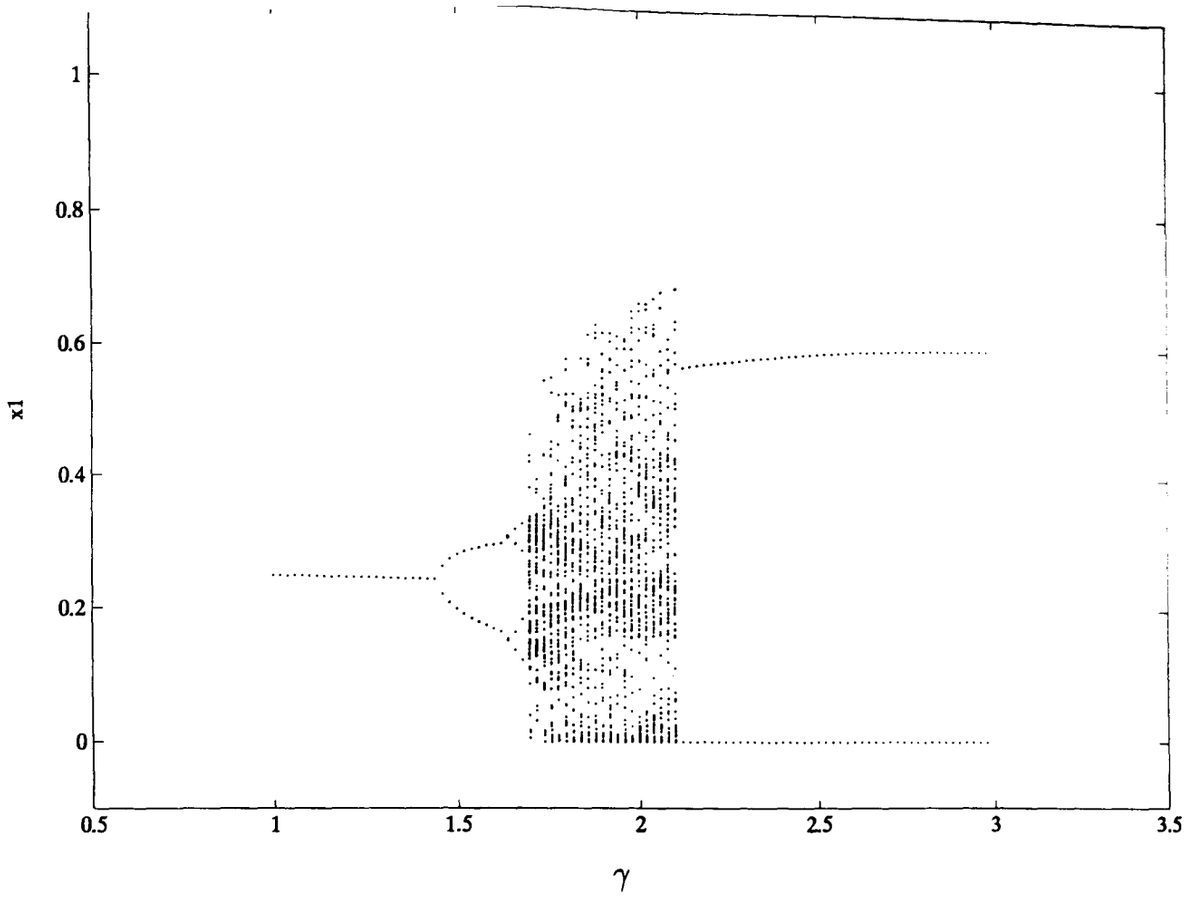
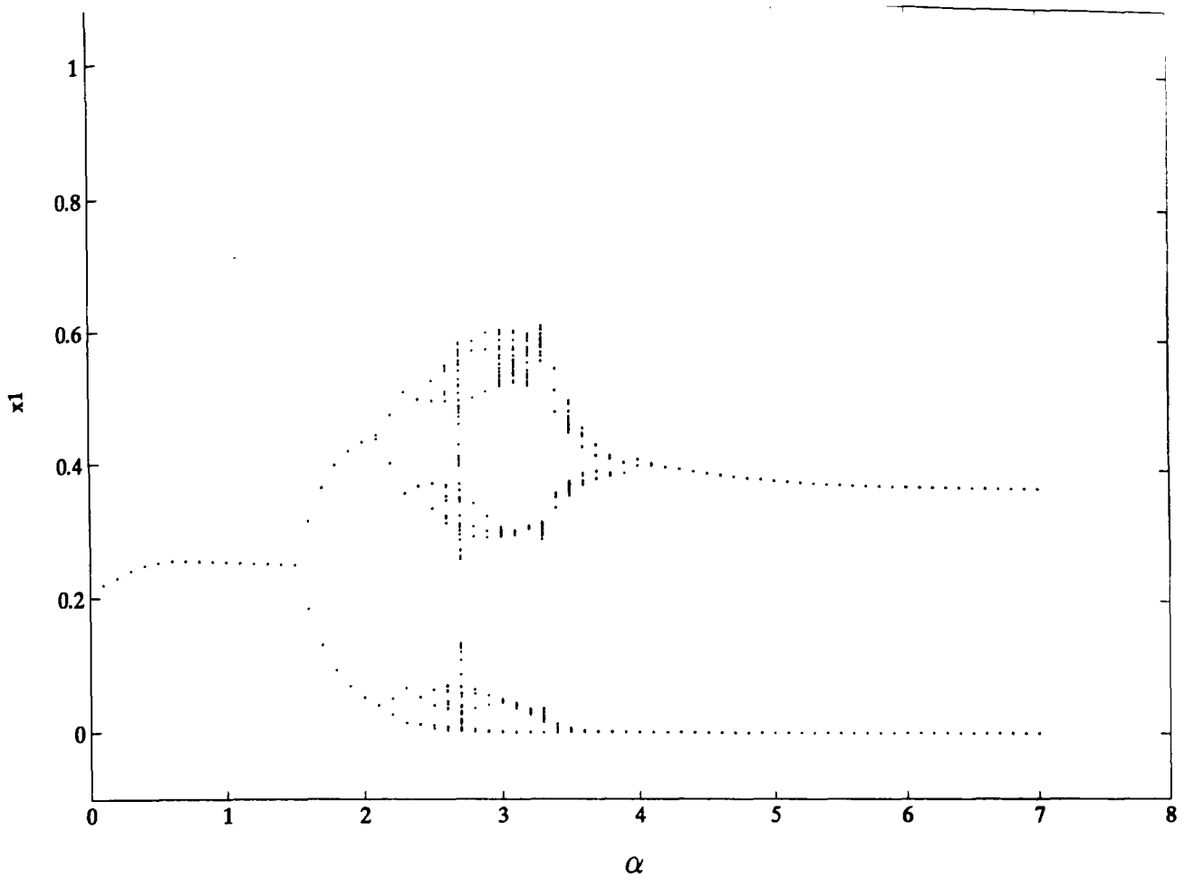
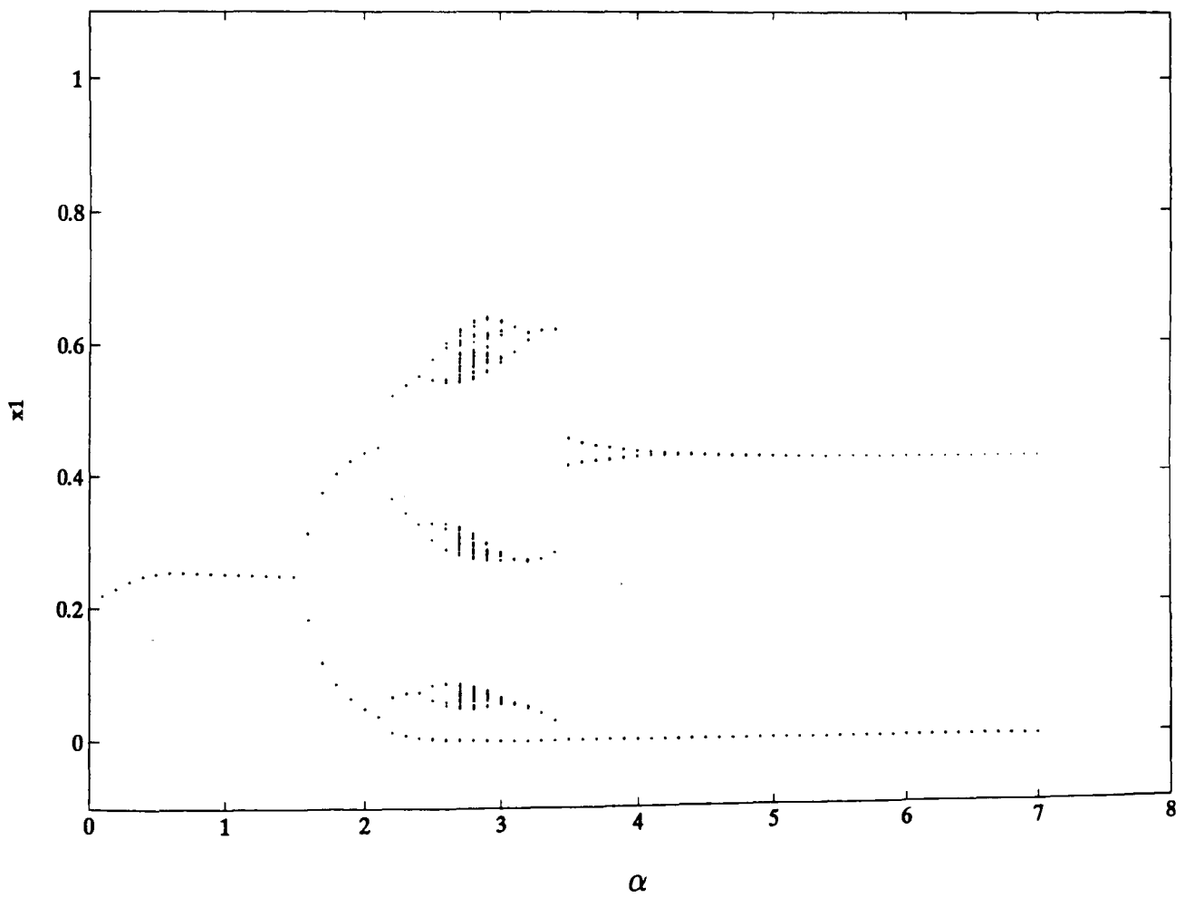


Figure 4. Bifurcation diagram of the unconstrained or singly constrained gravity model for γ, with $\beta = 1.5$, $\mu = 4.5$, $\alpha = 1.0$.

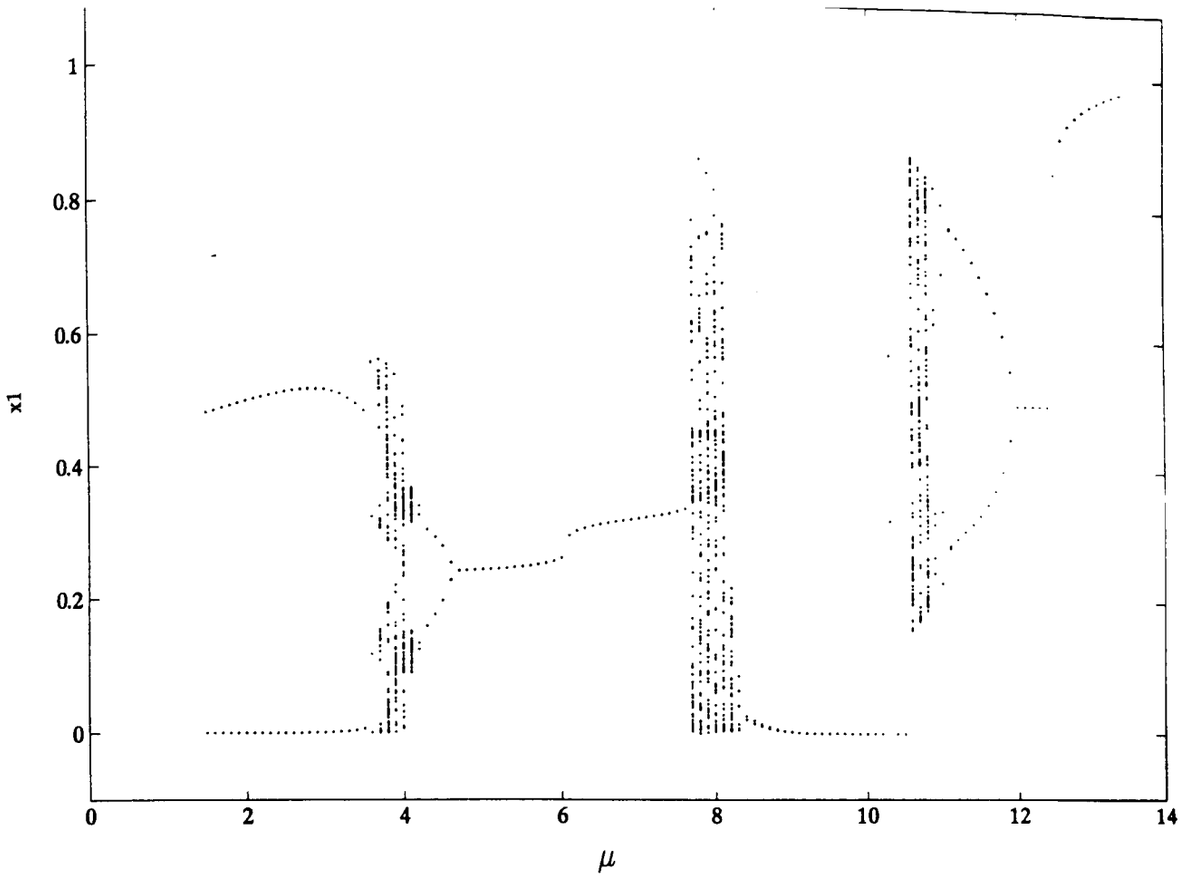


(a)

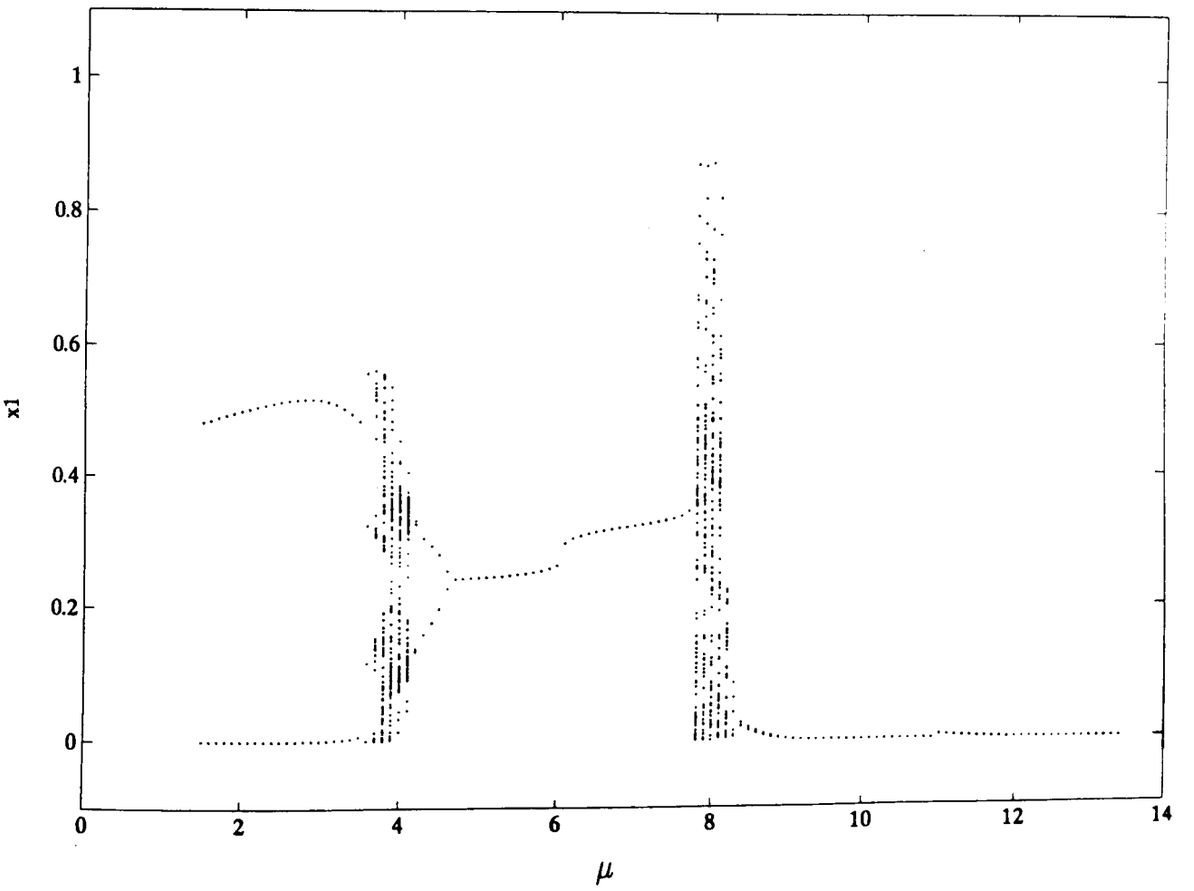


(b)

Figure 5. Bifurcation diagram of the unconstrained or singly constrained gravity model for α, with $\beta = 1.5$, $\mu = 4.5$, $\gamma = 0.75$. (a) Initial conditions are the same for all values of α; (b) initial conditions are the final states of the previous step of α.



(a)



(b)

Figure 6. Bifurcation diagram of the unconstrained or singly constrained gravity model for μ, with $\beta = 1.5$, $\alpha = 1.0$, $\gamma = 1.5$. (a) Initial conditions are the same for all values of μ; (b) initial conditions are the final states of the previous step of μ.

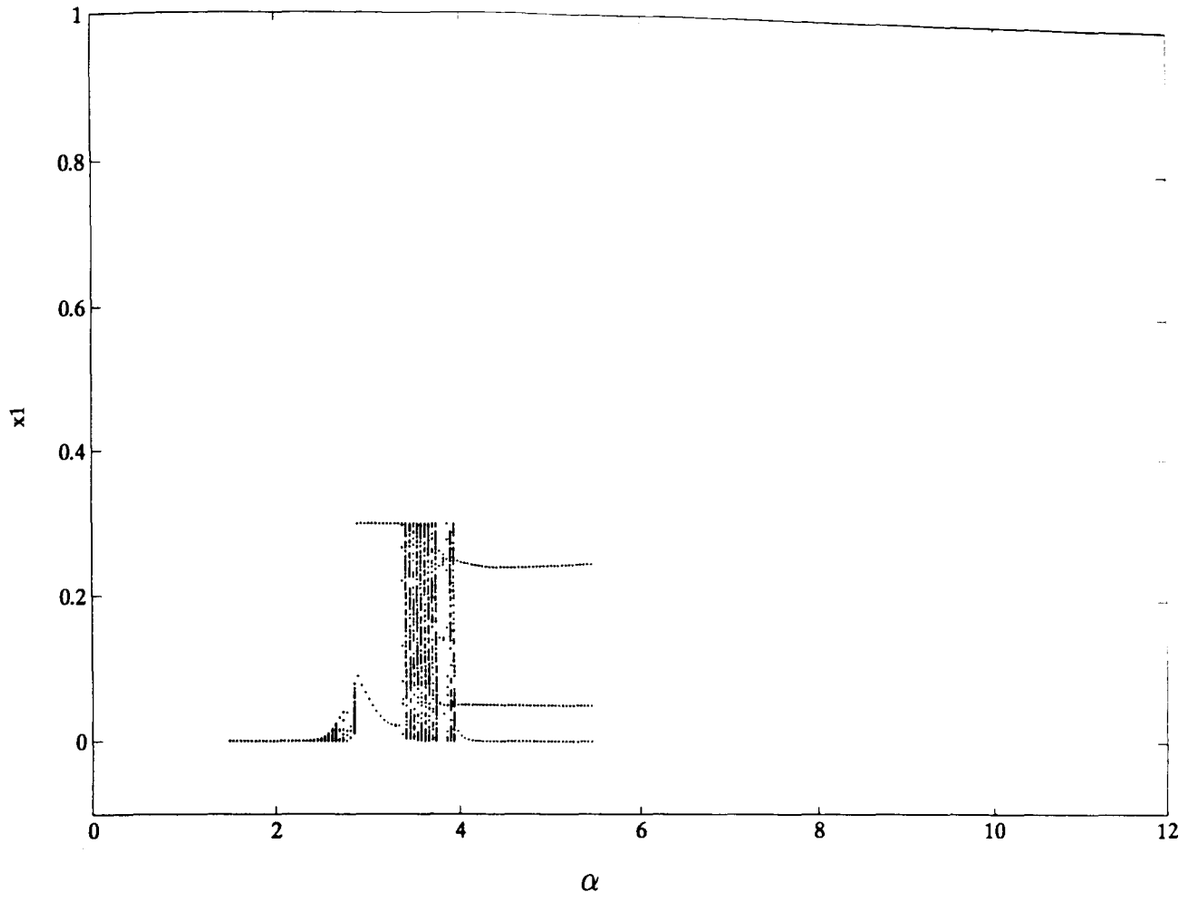


Figure 7. Bifurcation diagram of the doubly constrained gravity model for α, with $\beta = 1.0$, $\mu = 7.5$, $\gamma = 1.5$.

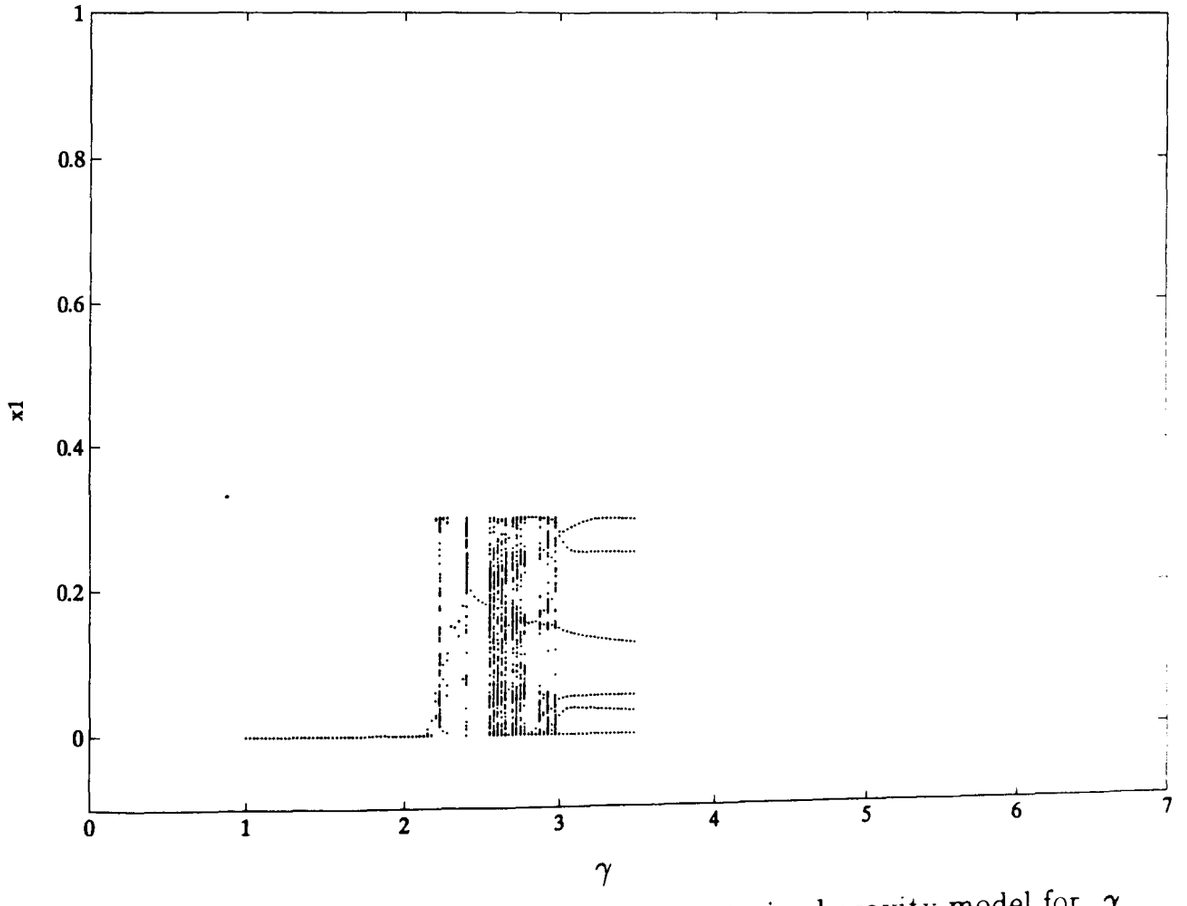


Figure 8. Bifurcation diagram of the doubly constrained gravity model for γ, with $\beta = 1.0$, $\mu = 7.5$, $\alpha = 1.5$.

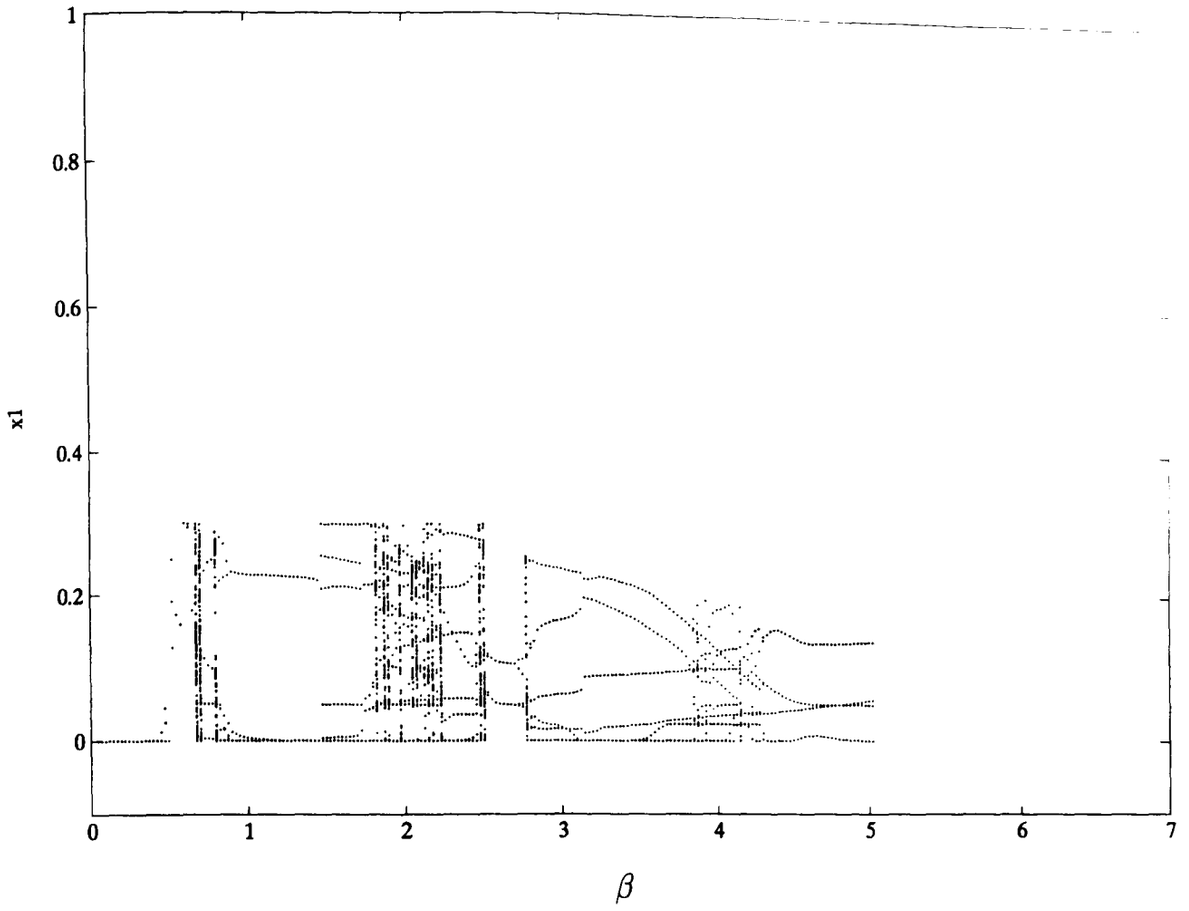


Figure 9. Bifurcation diagram of the doubly constrained gravity model for β, with $\mu = 7.5$, $\alpha = 2.5$, $\gamma = 2.5$.

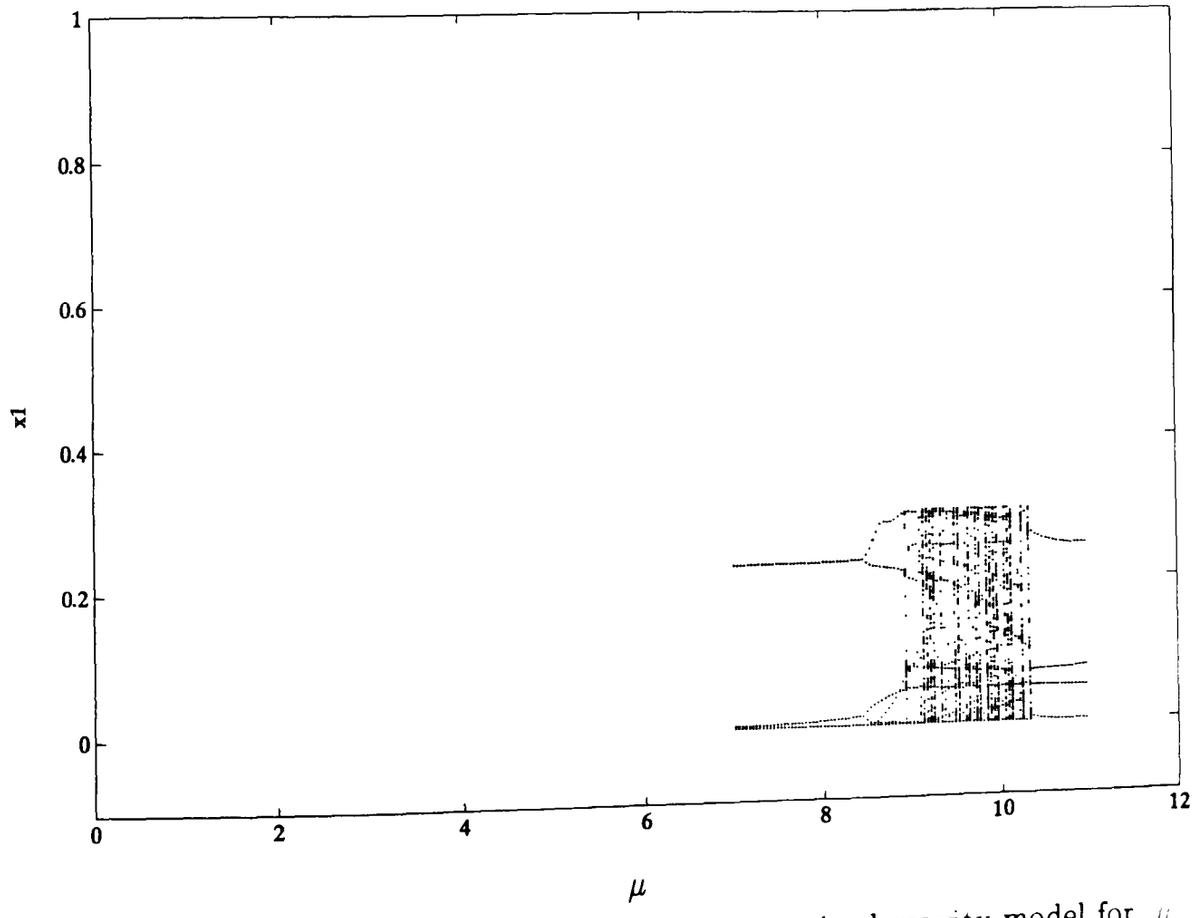


Figure 10. Bifurcation diagram of the doubly constrained gravity model for μ, with $\beta = 1.0$, $\alpha = 2.5$, $\gamma = 2.5$.

PUBLISHED AND CONFERENCE PAPERS

Two papers were written based on the research of the thesis during the term of this study. The first paper was published (Jarrett and Zhang, 1993) and the second paper was presented in the 26th Annual Conference of Universities Transport Study Group (Zhang, 1994). Copies of the two papers can be found at the end of this thesis.

The Dynamic Behaviour of Road Traffic Flow:

Stability or Chaos?

David Jarrett and Zhang Xiaoyan

Abstract

This paper is a report on work in progress on a project concerned with models of road traffic flow. Results for two such models are described and illustrated. One model is the classical car-following model. A number of numerical simulations were carried out, but no evidence of chaos was found. The other model concerns trip distribution. Here a dynamic formulation of the model results in some solutions which appear chaotic, and evidence of a period-doubling sequence of bifurcations is found.

Introduction

The motion of road traffic on a road network or on a single link of a network (a stretch of road between junctions) can be considered as a dynamic system. At a microscopic level, the system can be described in terms of variables such as the position and velocity of each vehicle. At a more macroscopic level, important variables include the total number of trips between two zones, the rate of traffic flow (the number of vehicles per unit time passing a fixed point), traffic density (the number of vehicles per lane and kilometre) and average speed. Dynamic models of road traffic flow describe how these variables change with time, possibly in response to external demands. These might be expressed, for instance, as the need for certain numbers of individuals or goods to be in particular places at particular times. The concepts of traffic equilibrium and its stability receive much attention; they are important both for understanding the behaviour of road traffic, and in traffic management and planning. Equilibrium and stability are desirable objectives for road traffic flow but are not always achieved. In dense traffic, where drivers follow each other very closely, small disturbances like the acceleration or deceleration of one vehicle might be preserved or amplified along the line of vehicles or over time, suggesting that there can be sensitive dependence on initial conditions. These phenomena can raise problems in traffic management, and can even result in accidents.

To the casual observer, road traffic flow appears inherently stochastic. However, many theoretical models of traffic flow are deterministic. Thus, it is appropriate to investigate these models and observed traffic flow for chaotic behaviour. This paper is a report on work in progress on a project concerned with nonlinear dynamic models of road traffic flow. One model being investigated is the car-following model, which describes the microscopic behaviour of congested traffic moving along a link. Two other recent studies investigated whether there is chaos in this model, but reached apparently different conclusions. This model and its solutions are considered in the next section. Macroscopic models being investigated include traffic assignment models and trip distribution models. Traffic assignment models are multi-dimensional discrete-time systems, in which the flows on the routes in the road network are iterated. In each iteration, the flows are diverted to 'cheaper routes' as perceived by drivers. Trip-distribution models are similar in form to assignment models, with the observables being the number of trips between given origin and destination zones. These models are considered later in the paper.

Computer systems are already widely used for the control of traffic signals, to move traffic efficiently through urban areas. Currently under development are systems for automated route guidance, and artificial intelligence systems in cars. One important application of this project is to investigate whether the introduction of such systems into the driving process, with the intention of improving traffic quality and efficiency, can lead to instability, or even chaos, in the motion of individual vehicles or the distribution of traffic over the system.

The Car-following Model

The motion of a line of vehicles on a crowded road link without overtaking is described by the car-following model [Leut88: Wilh73]. This model is based on the assumption that a driver responds to the motion of the vehicle immediately in front. In the simplest model, the acceleration of the following car is assumed to be proportional to the difference between its speed and that of the car in front; this model is linear. More complex models are nonlinear and allow the acceleration of the following car to depend both on its own speed and on the relative spacing of the two cars. In all cases, a time delay is built into the equations. The driver does not react immediately to changes in relative speed or spacing.

Consider a line of cars numbered from 1 (the leading car) to N (the last car) (Figure 1). Let $x_n(t)$ denote the position of car n at time t. Then the equations of the model are

$$\ddot{x}_n(t) = \alpha[\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)] \qquad n = 2, 3, \cdots, N$$

Here, α is a function of the current speed of car n, and its distance from car $n-1$ at the time $(t-\tau)$

$$\alpha = c \frac{(\dot{x}_n(t))^m}{[x_{n-1}(t-\tau) - x_n(t-\tau)]^l}$$

In these equations, dots denote time derivatives. τ is a constant, representing the reaction time of the driver of the following car. c is a positive parameter; m and l are nonnegative parameters, not necessarily integers. The linear model corresponds to α constant, with $m = l = 0$. In general, α is called the *sensitivity*.

Many stability studies of traffic flow on a road link are based on this model. Most recently, there have been two studies about possible chaotic behaviour in the car-following model. Disbro and Frame [Disb90] claim that chaos can definitely occur in the car-following process. Their conclusion is based on showing that the first Lyapunov exponent is positive for certain values of the parameters, but they give no plots or other evidence of a strange attractor. On the other hand, in an exploratory study Kirby and Smith [Kirb91] found no evidence of chaos in car following.

From a mathematical point of view, the car-following equations have a number of interesting features. Firstly and most importantly, they are a system of delay-differential equations and therefore have an infinite-dimensional phase space. Initial conditions must be specified as functions $x_n(t)$ over the interval $-\tau \leq t \leq 0$. The mathematical analysis of such equations is difficult. However, finite-dimensional attractors exist for systems of delay-differential equations; some such systems are believed to possess finite-dimensional strange attractors (see [Farm82]). Secondly, for given initial conditions, solutions do not necessarily exist for all $t \geq 0$; for $l > 0$ the equations have a singularity where $x_{n-1}(t-\tau) = x_n(t-\tau)$, corresponding to a collision between vehicles $n-1$ and n at time $(t-\tau)$. Such collisions can occur quite easily. Similarly, for m noninteger the equations make no sense if the speed of any vehicle becomes negative. Thirdly, where the solutions $x_n(t)$ exist for all positive t, they in general are unbounded: the cars eventually reach any given point on the road. It is also possible for the speed of one or more vehicles to increase without limit.

Note that the motion of each car is influenced directly by only the car immediately in front; the motion of the first car is taken as given. The equations

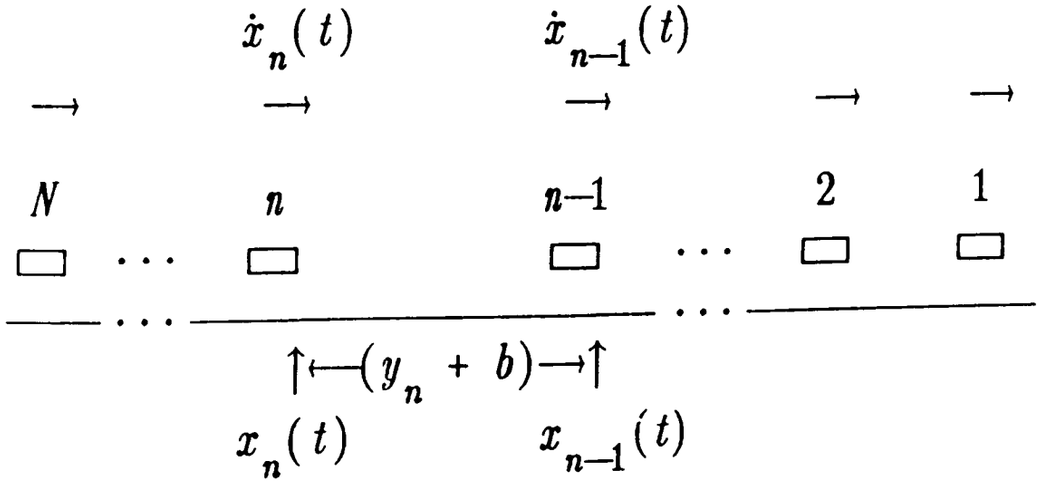


Figure 1. The car-following model.

can therefore be solved numerically one by one. However, progress can be made analytically by writing the equation for car n as

$$\frac{\ddot{x}_n(t)}{\dot{x}_n(t)^m} = c\,\frac{\dot{x}_{n-1}(t-\tau) - \dot{x}_n(t-\tau)}{\left[x_{n-1}(t-\tau) - x_n(t-\tau)\right]^l}$$

These equations can be integrated once (see [Leut88], p. 140).

Further analysis of the model is most easily carried out by re-expressing the model in terms of relative quantities so the solutions can remain bounded. Denote the spacing between adjacent cars by

$$y_n(t) = x_{n-1}(t) - x_n(t) - b \qquad n = 2, 3, \cdots, N$$

where b is interpreted as the minimum headway (see Figure 1). Then

$$\dot{y}_n(t) = \dot{x}_{n-1}(t) - \dot{x}_n(t)$$
$$\ddot{y}_n(t) = \ddot{x}_{n-1}(t) - \ddot{x}_n(t)$$

Let $x_1 \equiv x_1(t)$, $y_n \equiv y_n(t)$, $y_n^\tau \equiv y_n(t-\tau)$. Then the model becomes

$$\ddot{y}_2 = \ddot{x}_1 - c(\dot{x}_1 - \dot{y}_2)^m \frac{\dot{y}_2^\tau}{(y_2^\tau + b)^l}$$

$$\ddot{y}_n = c(\dot{x}_1 - \dot{y}_2 - \cdots - \dot{y}_{n-1})^m \frac{\dot{y}_{n-1}^\tau}{(y_{n-1}^\tau + b)^l}$$
$$- c(\dot{x}_1 - \dot{y}_2 - \cdots - \dot{y}_n)^m \frac{\dot{y}_n^\tau}{(y_n^\tau + b)^l} \qquad n = 3, 4, \cdots, N$$

These equations are solved numerically using a Runge-Kutta algorithm, modified for dealing with the delay time in the equations. The motion of a line of cars can be simulated, with the movement of the first car being treated as an input, that is

$$x_1(t) = ut + \frac{a}{\omega^2}\sin\omega t$$

Thus, the speed of the first car fluctuates about a constant, u. For $a = 0$ the equations are autonomous; for $a \neq 0$ there is a sinusoidal forcing term. This model has been simulated for selected combinations of values of the parameters m and l, for both the autonomous and the forced model.

If there is no forcing term, then typically there are stable equilibria, where the relative speed is zero and the relative spacing is constant. The limiting relative spacing depends on the initial conditions. There is a continuum of fixed points — note from the equations that if the relative speed becomes zero then the relative spacing remains constant. If the sensitivity, as determined by the parameter, c, is small, then most solutions converge to an equilibrium without oscillation. However, unbounded solutions can exist if the initial conditions are far from equilibrium. As the sensitivity increases, oscillations begin to occur.

For initial conditions close to an equilibrium, these oscillations damp down as the solution converges. If the initial conditions are further from equilibrium, then the solutions break down: they oscillate with increasing amplitude until the speed of the last car becomes negative. The basins of attraction of the equilibria become smaller as the sensitivity increases. In between the two cases, a periodic attractor (limit cycle) exists in the case $m = 0$ and $l = 2$ for a small range of values of the parameter, c. A solution converging to this attractor is shown in Figure 2. The period is approximately 4τ. This appears to be independent of the initial conditions, the value of c and the number of cars — only the amplitude of the oscillations depends on these factors. The basin of attraction of this periodic attractor is very small. A small change in the initial conditions results in a long transient, but the solution eventually tends to an equilibrium or breaks down.

When the forcing term is introduced, the solutions behave in a very similar way to those of the autonomous equations. The equilibria are replaced by stable periodic solutions with the same period as the forcing term. The amplitudes of these solutions depend on the initial conditions, sensitivity and the position of the car in the line. For initial conditions not close to a stable solution, the solutions diverge and break down, with long transients. Again, the basins of attraction of the periodic attractors become smaller as the sensitivity is increased. Where the unforced equations have a periodic attractor, the forced equations have a more complex behaviour. This is illustrated in Figure 3. However, this

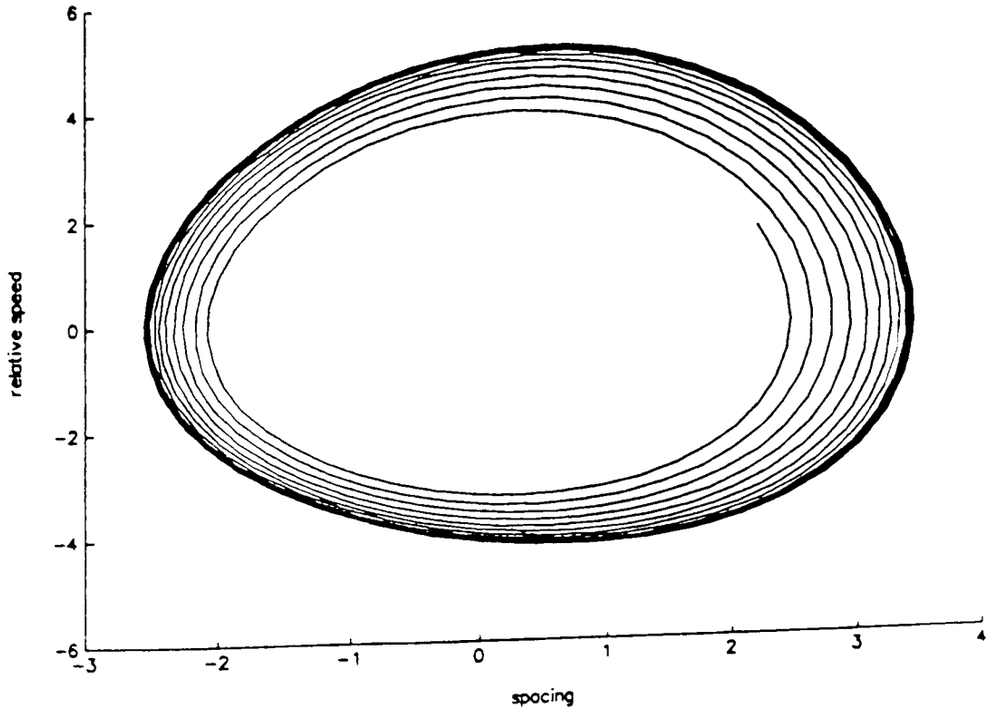
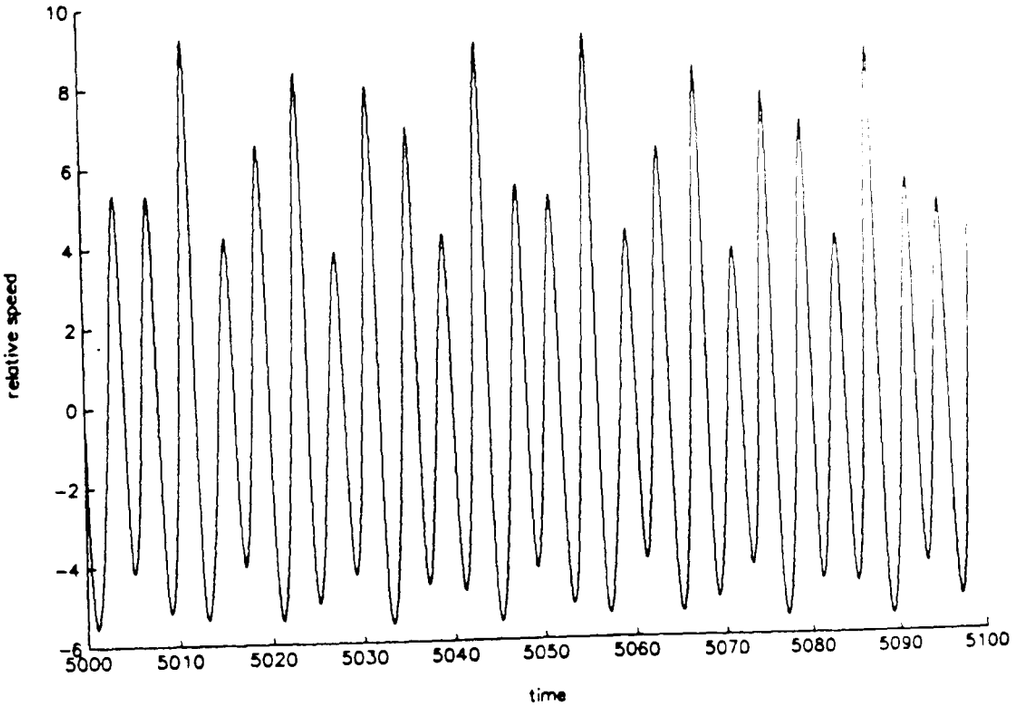
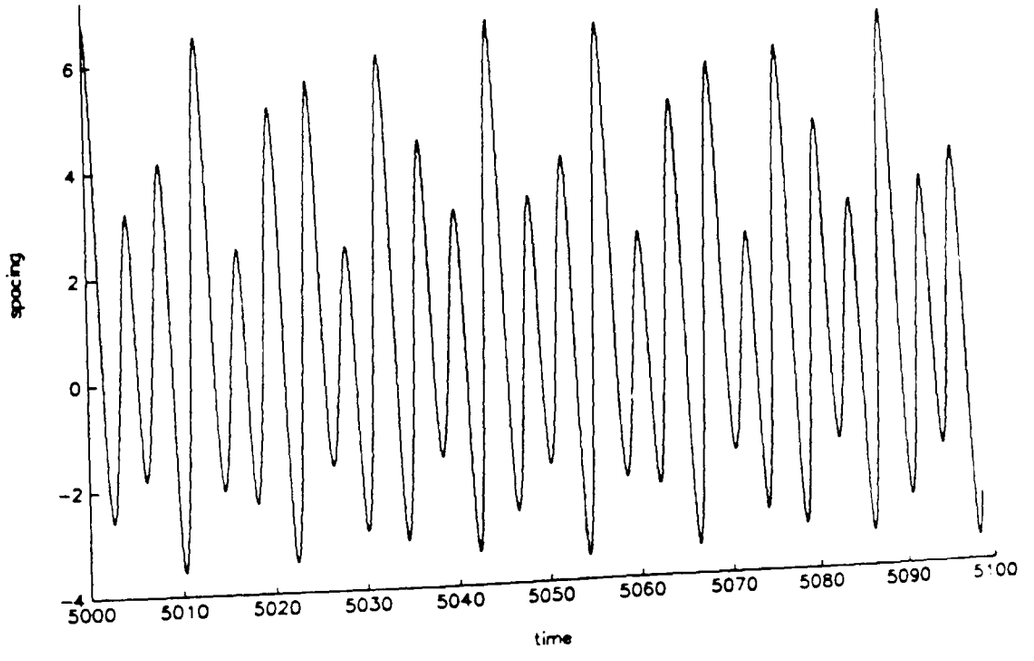


Figure 2. Unforced car-following model, periodic attractor, with $N = 4$, $m = 0$, $l = 2$, $\tau = 1$, $c = 155.3$ and $u = 15$. Relative speed and spacing of third and fourth cars.

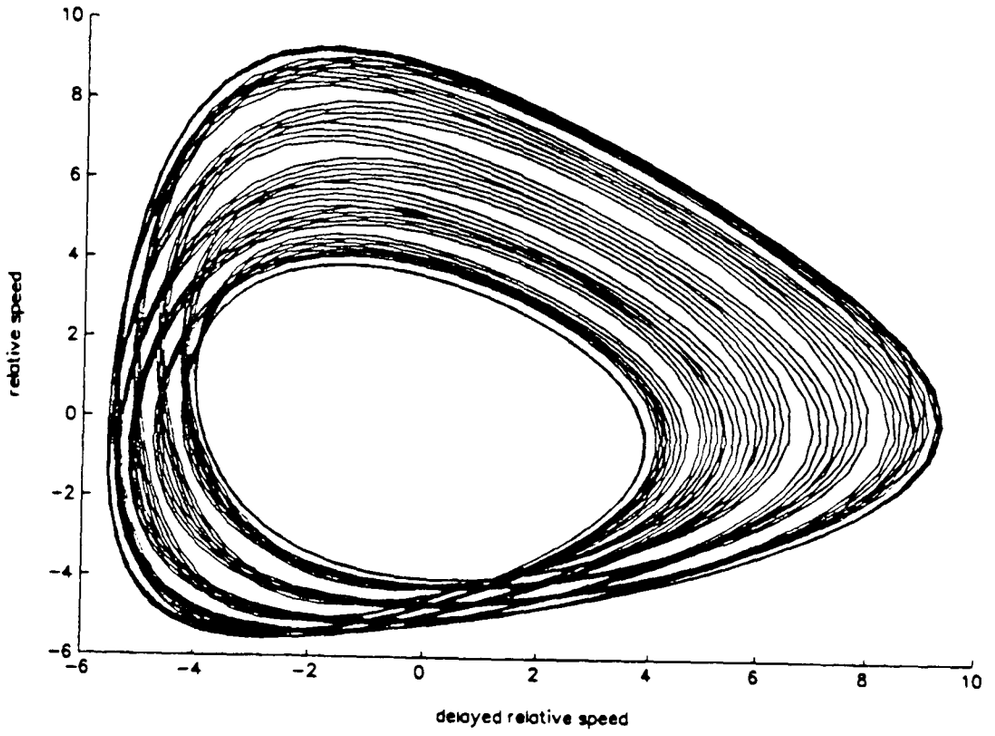


(a)



(b)

Figure 3. Forced car-following model, quasi-periodic solution, with $N = 4$; $m = 0$, $l = 2$, $\tau = 1$, $c = 165.3$; $u = 15$, $\omega = 1$, $a = 0.5$. (a) Relative speed; (b) relative spacing; (c) relative speed against delayed relative speed (delay = reaction time) of the third car to the fourth car.



(c)

Figure 3. (*Continued.*).

solution is not chaotic; it appears to be quasi-periodic, or possibly periodic with a very long period.

When the solutions oscillate, a small deviation in the initial conditions tends to be amplified along the line of cars, although it can damp down with time. Thus, any breakdown in the equations occurs only with the last car in the line. Removing car N averts the breakdown. However, a similar situation is bound to occur for car $N-1$ at a larger deviation from the equilibrium or an increased sensitivity. Thus, it is felt that the number of cars has little effect on the qualitative form of the solution.

In none of the investigations carried out so far has any evidence of chaos in the car-following model been found. Of course, no general conclusion can be reached from a finite number of numerical experiments. However, chaos does not appear to occur typically in the model. From a practical point of view, the existence of solutions leading to collisions is probably more important than the existence of chaotic solutions, and indicates the need for care if the car-following equations are used in any method for the automatic control of cars.

There is another study of chaos in road traffic flow by Kühne [Kuhn91], based on a fluid approximation model. In contrast to the car-following model, which treats vehicles individually, the fluid model describes traffic behaviour by macroscopic quantities such as flow rate, traffic density and average speed. Kühne

investigates a truncation of the fluid equations; he finds a chaotic attractor and computes the first Lyapunov exponent. The present authors have not yet investigated this model.

The Trip Distribution Model

The aim of the trip distribution model is to determine the number of trips between each pair of zones given the number of trips originating and terminating in each zone. One of the most widely used models is the gravity model, which assumes that the number of trips between zones depends on the number produced at and attracted to each zone, and on the travel cost between zones. Most formulations of this model are static, where the travel costs are assumed to be independent of the number of trips. Dendrinos and Sonis [Dend90] give a dynamic formulation in which the travel costs are a function of the number of trips between zones. At each stage of the iteration, the number of trips is generated using the costs associated with the trips of the previous stage. Dendrinos and Sonis suggest that this iteration can be chaotic, but they do not specify ranges of parameters for which this might be the case.

The dynamic trip distribution model takes the form

$$x_{ij}(t+1) = k(t) f\big(c_{ij}(t)\big)$$

where $x_{ij}(t)$ is the relative number of trips from zone i to zone j, normalised so that $\sum\sum x_{ij}(t) = 1$, and $c_{ij}(t)$ is the travel cost from zone i to zone j, given the trips $x_{ij}(t)$. Dendrinos and Sonis suggest taking

$$c_{ij}(t) = c_{ij}^0 \left[1 + \alpha\left(\frac{x_{ij}(t)}{z_{ij}}\right)^{\gamma}\right]$$

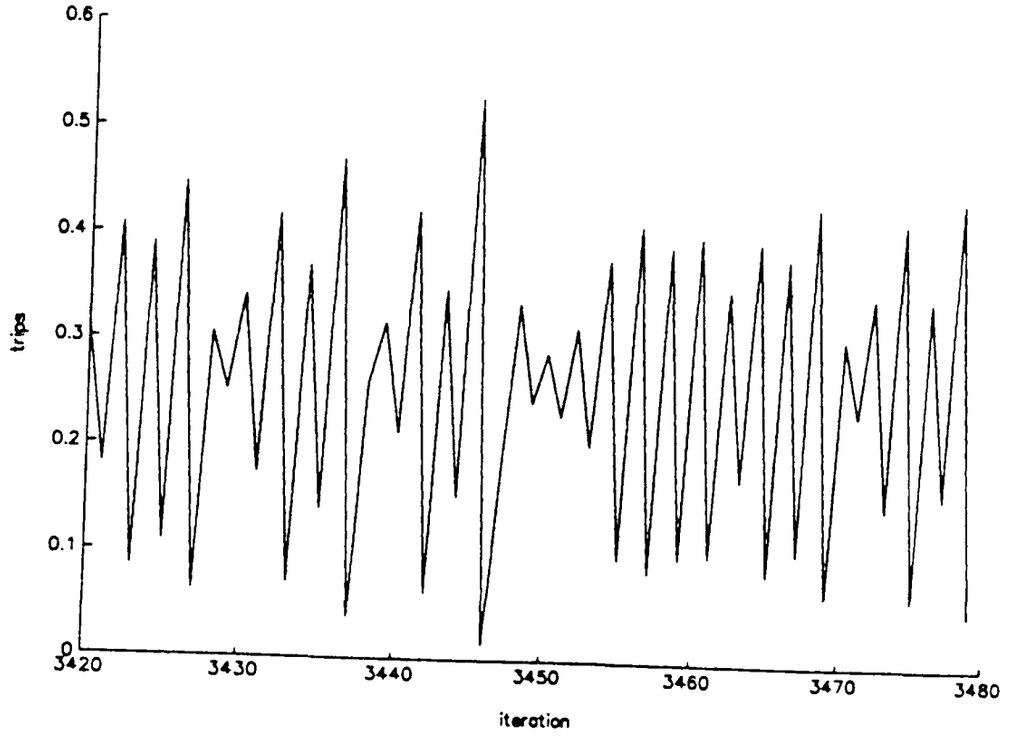
where c_{ij}^0 is the uncongested travel cost, z_{ij} is the relative capacity and α and γ are constants.

$f(c_{ij})$ is a function which relates the number of trips to the travel costs. Three types of function have been suggested (see [Ortu90]):

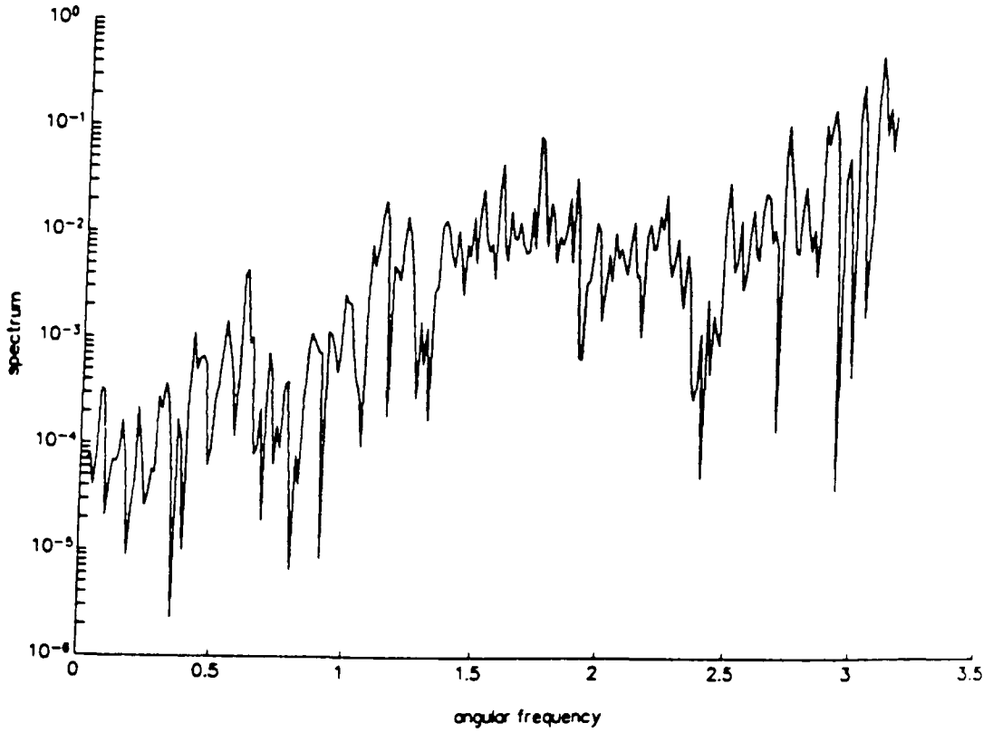
$$\begin{array}{lllll} \text{(a)} & f(c_{ij}) & = & \exp(-\beta c_{ij}) & \text{(exponential)} \\ \text{(b)} & f(c_{ij}) & = & c_{ij}^{-n} & \text{(power)} \\ \text{(c)} & f(c_{ij}) & = & c_{ij}^{n} \exp(-\beta c_{ij}) & \text{(combined)} \end{array}$$

where β and n are positive constants. For (a) and (b) the number of trips is a decreasing function of cost, while in (c) the number of trips first increases and then declines as cost increases.

The model as formulated above is unconstrained: it cannot guarantee that the number of trips originating from or terminating at a given zone has a value



(a)



(b)

Figure 4. Trip distribution model, chaotic solution with two origins and two destinations, $\alpha = 1$, $\gamma = 1$; $n = 8$, $\beta = 3.25$. (a) Number of trips from origin 1 to destination 2; (b) power spectrum.

which is predetermined. In fact, the numbers become totally different from the starting values after a few iterations. Constraints can be incorporated but are not investigated here. This model has been investigated with two or three origin and destination zones, using each of the three forms of cost functions. In most cases there is a stable fixed point or periodic orbit. However, for the combined cost function, evidence of a period-doubling sequence of bifurcations has been found. If α and γ are fixed at 1, then for appropriate values of n the period changes from 2 to 4 to 8 to 16 as β is gradually increased. Eventually, the sequence becomes irregular and appears chaotic, with a continuous power spectrum. Figure 4 shows such a solution for the three-dimensional model with two origins and two destinations. Lyapunov exponents for this solution were calculated using the algorithm of Eckmann and Ruelle [Eckm85], and were found to be approximately 0.2, -0.02 and -0.7, respectively. A bifurcation diagram for this model is shown in Figure 5. The apparent discontinuity seems to be because there are two or more periodic attractors for some parameter values.

Related models concern traffic assignment. Given the trips $\left(x_{ij}(t)\right)$, the models attempt to estimate the flow on each link of a road network. The assignment of flows to different links is a dynamic process; in each iteration the flows are diverted to 'cheaper routes' as perceived by drivers. When the flow on each link tends to a constant value, it is said to converge to an equilibrium state. Horowitz [Horo84] pointed out that even in a two-route system this equilibrium might not be stable, that is, the equilibrium might not be reached or approached from arbitrary initial conditions. The authors are currently investigating these models and hope to report on them elsewhere.

Conclusion

In this paper two different traffic models were investigated for the presence of chaotic solutions. No evidence of chaos was found in the car-following model, although other authors have indicated that chaos can be found in related models. However, chaotic solutions were found in a dynamic trip-distribution model. It is clear that other traffic models might have chaotic solutions, and the authors hope to report on these in a later paper.

Acknowledgments. The authors thank Professor C.C. Wright for comments on an earlier draft of this paper. The second author is grateful to Middlesex University for financial support.

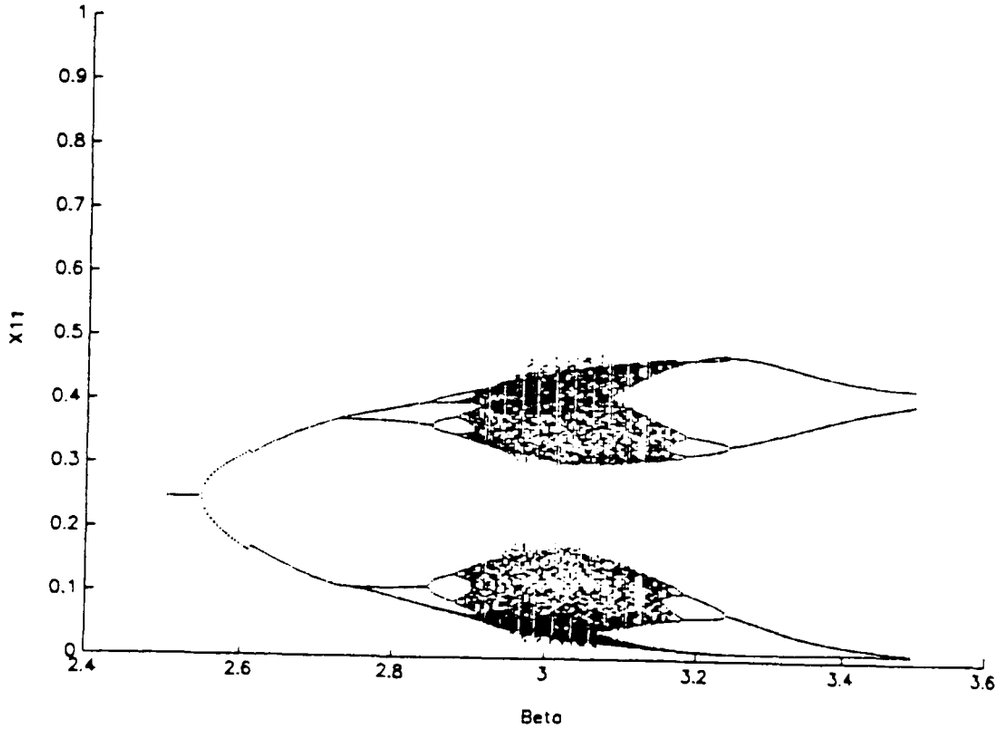


Figure 5. Trip distribution model, bifurcation diagram with two origins and two destinations; $\alpha = 1$, $\gamma = 1$; $n = 7$. The vertical axis is the number of trips from origin 1 to destination 1.

REFERENCES

[Dend90]
Dendrinos, D.S., and Sonis, M., *Chaos and Social-Spatial Dynamics*, New York: Springer-Verlag, 1990.

[Disb90]
Disbro, J.E., and Frame, M., Traffic flow theory and chaotic behaviour, *Transportation Research Record*, Vol. 1225, pp. 109–115, 1990.

[Eckm85]
Eckmann, J-P., and Ruelle, D., Ergodic theory of chaos and strange attractors, *Rev. Modern Physics*, Vol. 57, No. 3, pp. 617–656, 1985.

[Farm82]
Farmer, J.D., Chaotic attractors of an infinite-dimensional dynamical system, *Physica*, Vol. 7D, pp. 366–393, 1982.

[Horo84]
Horowitz, J.L., The stability of stochastic equilibrium in a two-link transportation network, *Transportation Research*, Vol. 18B, pp. 13–28, 1984.

[Kirb91]
Kirby, H.R., and Smith, M.J., Can chaos theories have transport applications?, 23rd Annual Conf. of the Universities Transport Studies Group, University of Nottingham, 1991.

[Kuhn91]
Kühne, R.D., Chaotic behaviour of traffic flow, Preprint, 1991.

[Leut88]
Leutzbach, W., *Introduction to the Theory of Traffic Flow*, Berlin: Springer-Verlag, 1988.

[Ortu90]
Ortúzar, J.deD., and Willumsen, L.G., *Modelling Transport*, Chichester, UK: Wiley, 1990.

[Wilh73]
Wilhelm, W.E., and Schmidt, J.W., Review of car following theory, *Transportation Engrng. Jour.* (ASCE), Vol. 99, TE4, pp. 923–933, 1973.

TRAFFIC DYNAMICS: STUDIES IN THE GRAVITY MODEL AND THE TRIP ASSIGNMENT MODEL

Xiaoyan Zhang
Middlesex University

Abstract

The aim of this paper is to investigate the dynamic properties of non-linear dynamic models of road traffic flow. One model considered is the dynamic version of the gravity model proposed by Dendrinos and Sonis (1990) . In a previous paper (Jarrett and Zhang, 1993) this dynamic model was studied numerically and the period–doubling route to chaos was found. In this paper, a theoretical analysis is made of a one–dimensional model; Liapunov exponents and correlation dimensions are calculated for (chaotic) attractors found in the multidimensional gravity models and are found to be positive and fractal for chaotic attractors. Another traffic model studied here is the dynamic, logit–based trip assignment model, also suggested by Dendrinos and Sonis (1990). Stabilities of fixed points and period–two orbits are studied for a network composed of one O–D pair connected by two links.

1. INTRODUCTION

Dynamic traffic flow problems on a road network have been modelled and studied in different ways. One way is to consider explicitly the time–dependence of network characteristics, such as costs, and traffic flows. Dynamic link performance functions are used to consider queueing and congestion effects on links; the distribution of departure times of O–D trips depends on the temporal distribution of travel costs over the network so that the O–D flows vary with time. Another type of dynamic consideration is to model the process of adjustments of flow pattern in a network from one time instant to another for a given O–D matrix. This second type of study has received little attention and is considered here. Horowitz (1984) studied the stability of a stochastic equilibrium in a discrete–time assignment model for a network of one O–D pair connected by two links. Smith (1984) proposed a continuous–time adjustment mechanism modelled by a set of ordinary differential equations. The equilibrium of this dynamic system coincides with the Wardrop user equilibrium. Using a method due to Liapunov he was able to prove that the equilibrium is stable if the cost–flow function is monotonic and smooth.

The time evolution of a dynamical system is normally modelled by differential or difference equations. In practical applications these equations will usually be non–linear. Many dynamical systems exhibit a start–up transient, after which the motion settles down towards some form of steady–state behaviour. Motions from neighbouring initial values tend to converge towards stable attracting solutions called *attractors*. There are basically three types of attractors: point attractors, periodic attractors, and chaotic attractors of the kind which have been

discovered in the last twenty to thirty years. Stable or attracting limit sets (a limit set is a set of points in the phase space which a trajectory repeatedly visits) are of special interest since a non–stable limit set cannot be observed in real systems and simulations. Studies of a dynamical system often involve identifying possible attractors and analyzing how these change with value of parameters.

Chaos is a kind of irregular behaviour found in deterministic systems. It is not yet fully understood. Therefore, it is of theoretical interest to find out if a model possesses chaos and to study chaotic attractors. As far as traffic studies are concerned, however, knowledge of the range of parameters for different types of behaviour in a model is clearly important. Values of parameters can be chosen to avoid unwanted behaviour such as instabilities, oscillations, and chaos; traffic flow can be monitored to achieve a stable equilibrium or even a better one if there are more than one.

In this paper, the time variation of traffic flow is modelled by deterministic, dynamical systems expressed as difference equations. Traffic dynamics in trip distribution and trip assignment are studied, respectively based on a dynamic gravity model and a dynamic logit–based trip assignment model, both suggested by Dendrinos and Sonis (1990). Theoretical analyses are made of one–dimensional models and numerical study is employed to calculate Liapunov exponents and correlation dimensions for (chaotic) attractors found in multidimensional gravity model. The gravity model is investigated in the next section, followed by the study of the trip assignment model in the subsequent section. The paper is summarized in the last section.

2. THE GRAVITY MODEL

The aim of gravity models is to estimate the number of trips between each O–D pair based on travel costs between zones and the total number of trips from and/or to each zone, or the total number of trips in the whole area under consideration. A family of gravity models in terms of relative quantities can be written in a general form

$$t_{ij} = \psi f(c_{ij}), \quad t_{ij} \in [0,1], \tag{1}$$

$$i=1, 2, \ldots, I, \quad j=1, 2, \ldots, J.$$

Here t_{ij} is the relative number of trips from zone i to zone j, c_{ij} is the corresponding travel cost, I and J are the number of origins and destinations respectively, $f(c_{ij})$ is called the *deterrence function* which relates the number of trips to the travel costs, and ψ is an appropriate normalizing factor. Three types of deterrence function have been suggested (Ortúzar & Willumsen, 1990): (a) exponential function, (b) power function, and (c) combined function. They can be written as:

$$f(c_{ij}) = c_{ij}^{\mu} \exp(-\beta c_{ij}),$$

where μ and β are constants which determine the form of the function. When $\mu=0$ and $\beta>0$, f is an exponential function; when $\mu<0$ and $\beta=0$ it is a power function; and when $\mu>0$ and $\beta>0$ it is a combined function. In the former two forms the number of trips is a decreasing function of cost, while in the

third the number of trips first increases and then declines as cost increases, depending on the relative magnitude of μ and β. The factor ψ in (1) is chosen so that one or two of the following constraints of an O–D matrix are satisfied

$$\text{(a)} \quad \Sigma_{ij} t_{ij} = 1, \tag{2a}$$

$$\text{(b)} \quad \Sigma_j t_{ij} = o_i, \qquad i=1, 2, \ldots, I, \tag{2b}$$

$$\text{(c)} \quad \Sigma_i t_{ij} = d_j, \qquad j=1, 2, \ldots, J, \tag{2c}$$

where o_i is the total (relative) number of trips originated from zone i, and d_j is the total (relative) number of trips terminated at zone j. A gravity model is unconstrained if only the total number of trips attracted to (or produced at) all zones (which has been normalized to 1 here) is known, or singly constrained if o_i's or d_j's are at hand, or doubly constrained if both o_i's and d_j's are given. The form of ψ for the three cases are as follows

(a) Unconstrained model. Only (2a) is satisfied and

$$\psi = \frac{1}{\Sigma_{ij} f(c_{ij})}$$

so that

$$t_{ij} = \frac{f(c_{ij})}{\Sigma_{ij} f(c_{ij})}.$$

(b) Singly constrained model. For an origin–constrained model (2b) is met and ψ is replaced by a set of constants:

$$a_i = o_i \frac{1}{\Sigma_j f(c_{ij})}, \qquad i = 1, 2, \ldots I,$$

so that

$$t_{ij} = o_i \frac{f(c_{ij})}{\Sigma_j f(c_{ij})}.$$

For a destination–constrained model (2c) is met and ψ is replaced by another set of constants:

$$b_j = d_j \frac{1}{\Sigma_i f(c_{ij})}, \qquad j = 1, 2, \ldots J,$$

so that

$$t_{ij} = d_j \frac{f(c_{ij})}{\Sigma_i f(c_{ij})}.$$

(c) Doubly constrained model. Both (2b) and (2c) are satisfied and ψ is replaced by two sets of constants, or the *balancing factors*:

$$a_i = o_i \frac{1}{\Sigma_j b_j f(c_{ij})}, \quad i = 1, 2, \ldots I,$$

$$b_j = d_j \frac{1}{\Sigma_i a_i f(c_{ij})}, \quad j = 1, 2, \ldots J,$$

and

$$t_{ij} = a_i b_j f(c_{ij}) .$$

Most gravity models in the literature are static and the travel costs are assumed to be independent of the number of trips. A dynamic model is needed to study the O–D flow dynamics in an area. Dendrinos and Sonis (1990) proposed an iterative version of the gravity model by assuming that the number of trips at each stage depends on the travel cost which, in turn, is a function of the number of trips at the previous stage, that is

$$t_{ij}(n+1) = \psi f(c_{ij}(n))$$

with

$$c_{ij}(n) \equiv c_i(t_i(n)) = c_{ij}^0 \left[1 + \alpha \left[\frac{t_{ij}(n)}{q_{ij}}\right]^\gamma\right],$$

where c_{ij}^0 is the uncongested travel cost, q_{ij} is the relative capacity, and α and γ are positive constants.

The dynamic unconstrained model is a map of the simplex

$$\Delta_{IJ} = \{[t_{ij}]: \ t_{ij} \geq 0, \ \Sigma_{i=1}^{I} \Sigma_{j=1}^{J} t_{ij} = 1 \} ;$$

into itself. The dimension of the phase space is IJ–1. The origin–constrained model is a map on

$$S_1 \times S_2 \times \ldots \times S_I$$

where

$$S_i = \{(t_{i1}, t_{i2}, \ldots, t_{iJ}): \ t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i \} , \quad i=1, 2, \ldots, I,$$

so that

$$S_1 \times S_2 \times \ldots \times S_I = \{[t_{ij}]: \ t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i, \ i=1, 2, \ldots, I \}$$

It can be seen that in the origin–constrained model t_{ij} depends only on the elements of the ith row of a trip matrix. Therefore this model can be considered as I independent equations equivalent to the unconstrained model, but on $S_i = \Delta_J$ and with a dimension of J–1. Similarly, the destination–constrained model is equivalent to J independent equations on Δ_I and with a dimension of I–1. The doubly–constrained model, however, is different. The phase space is

$$S = \{[t_{ij}]: \ t_{ij} \geq 0, \ \Sigma_j t_{ij} = o_i, \ i=1, 2, \ldots, I, \ \Sigma_i t_{ij} = d_j , \ j=1, 2, \ldots, J \}$$

$$= \{(t_{1\cdot}, t_{2\cdot}, \ldots, t_{I-1\,\cdot}) \in \Delta_J \times \Delta_J \times \ldots \times \Delta_J: \ \Sigma_i t_{ij} \leq d_j, \ j=1, 2, \ldots, J\}.$$

This phase space has a dimension of $(I-1)(J-1)$ and is more complicated than Δ_{IJ}. This paper will concentrate in the unconstrained and singly constrained models only. The doubly constrained model will be considered in a future paper.

The unconstrained and singly–constrained models can be written in the more general form, using single subscripts for simplicity,

$$t_i(n+1) = F_i(\boldsymbol{t}(n)) = \frac{f(c_i(n))}{\Sigma_j f(c_j(n))} \tag{3}$$

where

$$f(v) = v^{\mu} \exp(-\beta v)$$

$$c_i(n) = c_i(t_i(n)) = c_i^0 \left[1 + \alpha \left[\frac{t_i(n)}{q_i}\right]^{\gamma}\right]$$

$$i = 1, 2, \ldots, K.$$

When $K = IJ$ the equation represents an unconstrained gravity model; while when K equals I or J the equation represents one component of a singly constrained model.

The existence of a fixed point (in a discrete system an equilibrium is often called a fixed point) in (3) can be established by the Brouwer's fixed point theorem which can be stated as follows.

Brouwer's fixed point theorem Any continuous map $F: \Delta_K \to \Delta_K$ has at least one fixed point.

In (3), $f(v)$ and $c_i(n)$ are both continuous functions over the interval under consideration. Therefore, F_i is continuous and so, according to Brouwer's fixed point theorem, has at least one fixed point. A general analysis of the uniqueness and the stability of a fixed point in (3) is difficult for it is nonlinear. A one dimensional–model is analyzed in the next subsection; and higher dimensional–models are studied numerically in the following subsection.

2.1 Theoretical analysis — a one–dimensional model

A one dimensional model can occur where, for example, there are one origin and two destinations, or vice versa. The model is:

$$t_i(n+1) = F_i(\boldsymbol{t}(n)) = \frac{f(c_i(n))}{f(c_1(n)) + f(c_2(n))}, \quad i = 1, 2 \tag{4}$$

As mentioned above, this model has at least one fixed point. A fixed point t^e can be found by solving the nonlinear equation

$$t_i^e = F_i(t^e) .$$

The equation cannot be solved analytically unless F_i is such that the equation is

linear or quadratic.

It can be shown (See the Appendix) that if $f(c_i(t_i))$ is a monotonically decreasing function of t_i, then so is F_i, and the fixed point is unique. The derivative of f with respect to t_i is

$$\frac{\mathrm{d}}{\mathrm{d}t_i} f(c_i(t_i)) = f'(c_i)\, c_i'(t_i),$$

where

$$f'(v) = \mu\, v^{\mu-1}\exp(-\beta\, v) - \beta\, v^{\mu}\exp(-\beta\, v)$$

$$= f(v)\left(\frac{\mu}{v} - \beta\right),$$

$$c_i'(t_i) = c_i^0\, a\, \frac{\gamma}{q_i}\left(\frac{t_i}{q_i}\right)^{\gamma} \geq 0$$

The sign of $f'(v)$ depends on β and μ. When (a) $\beta=0$, $\mu<0$, and (b) $\mu=0$, $\beta>0$, $f'(v)$ is negative. And so the fixed point is unique. When $\mu>0$, and $\beta>0$, F_1 is not monotone, and the uniqueness cannot be assured.

A related problem of a fixed point is its stability. Although the fixed point cannot be obtained analytically in the more general case, the stability of it can be studied to some extent. A fixed point is asymptotically stable if the derivative at the point is within (–1, 1). Evaluating the derivatives at the fixed point t^e gives

$$\left|\frac{\mathrm{d}}{\mathrm{d}t_1} F_1(t^e)\right| = \left| t_1^e\, t_2^e \left(c_1'(t_1^e)\, \left[\frac{\mu}{c_1(t_1^e)} - \beta\right] \right.\right. \tag{5}$$

$$\left.\left. + c_2'(t_2^e)\, \left[\frac{\mu}{c_2(t_2^e)} - \beta\right] \right)\right|,$$

where

$$c_i'(t_i^e) = c_i^0\, a\, \frac{\gamma}{q_i}\left(\frac{t_i^e}{q_i}\right)^{\gamma}$$

Three types of deterrence function can be considered separately.

(a) When $\beta=0$ and $\mu<0$

$$\left|\frac{\mathrm{d}}{\mathrm{d}t_1} F_1(t^e)\right| = \left|\mu\gamma \left(\frac{c_1^0\ a(t_1^e/q_1)^{\gamma}}{c_1^0\ (1 + a(t_1^e/q_1)^{\gamma})}\, t_2^e + \frac{c_2^0\ a(t_2^e/q_2)^{\gamma}}{c_2^0\ (1 + a(t_2^e/q_2)^{\gamma})}\, t_1^e \right)\right|$$

$$\leq |\mu\gamma|\ (\ t_2^e + t_1^e\ ) = |\mu\gamma|$$

If $|\mu\gamma|<1$ then there must be

$$\left| \frac{\mathrm{d}}{\mathrm{d}t_1} F_1(t^{\mathrm{e}}) \right| < 1$$

and the fixed point is stable.

(b) When $\beta>0$ and $\mu=0$

$$\left| \frac{\mathrm{d}}{\mathrm{d}t_1} F_1(t^{\mathrm{e}}) \right| = \beta\, t_1^{\mathrm{e}}\, t_2^{\mathrm{e}} \left(c_1'(t_1^{\mathrm{e}}) + c_2'(t_2^{\mathrm{e}}) \right)$$

$$= \beta\, \alpha\, \gamma\, t_1^{\mathrm{e}}\, t_2^{\mathrm{e}} \left(\frac{c_1^0}{q_1} (t_1^{\mathrm{e}}/q_1)^{\gamma-1} + \frac{c_2^0}{q_2} (t_2^{\mathrm{e}}/q_2)^{\gamma-1} \right)$$

$$\leq 0.25\, \beta\, \alpha\, \gamma \left(\frac{c_1^0}{q_1} (t_1^{\mathrm{e}}/q_1)^{\gamma-1} + \frac{c_2^0}{q_2} (t_2^{\mathrm{e}}/q_2)^{\gamma-1} \right).$$

So if

$$\beta\, \alpha\, \gamma \left(\frac{c_1^0}{q_1} (t_1^{\mathrm{e}}/q_1)^{\gamma-1} + \frac{c_2^0}{q_2} (t_2^{\mathrm{e}}/q_2)^{\gamma-1} \right) < 1/0.25 = 4$$

the fixed point is stable.

(c) When $\beta>0$ $\mu>0$, the fixed point is not unique and no conclusion can be drawn about the stability.

The fixed points in the cases (a) and (b) lose stability as the parameters vary (μ decreases, β, α, and γ increase) such that the derivatives pass -1. Because the derivatives in both cases are negative, the fixed point bifurcates into a stable period–2 orbit through a flip bifurcation (see the Appendix). In the one–dimensional model with the exponential and the power deterrence functions fixed points and period–2 orbits are only possible steady states; trajectories starting from any initial conditions in the phase space approach one of the states.

It can be seen that for case (a) and (b), for given road conditions (described by c_i^0 and q_i), the stability of the fixed point depends on μ, β, α, and γ. The value of parameters reflects the inter–dependence of the number of trips and the travel costs. Higher values (or lower values for μ) mean a stronger dependence and can cause oscillations. Models of higher dimension and with a combined deterrence function are studied in the next subsection.

2.2 Numerical study — higher–dimensional models

Numerical study was made of (3) with various dimensions and with the three types of deterrence function in a previous paper (Jarrett and Zhang, 1993). Fixed points and period–2 orbits were found to be the typical behaviour in models with the exponential and power deterrence functions; a periodic–doubling route to chaos was found when the combined deterrence function ($\mu>0$ and $\beta>0$) is used. Chaotic attractors can be characterized by their Liapunov exponents and fractal dimensions, the former reflects the dynamical aspects and the latter the geometrical ones of an attractor. Calculations of these two quantities for

(chaotic) attractors found in (3) follow.

Calculation of Liapunov exponents

A Liapunov exponent is a measure of the average rate of change of small separations on an attractor. For a map

$$t_i(n+1) = F_i(t(n)),$$

Liapunov exponents are defined as

$$\lambda_i = \text{Log} \mid \sigma_i \mid,$$

where the σ_i are the eigenvalue of the following operator

$$\lim_{n\to\infty} \left(T_t^{n*} \; T_t^n \right)^{1/2n},$$

In this expression, T_t^{n*} stands for the adjoint of T_t^n, and T_t^n is the derivative of the nth iteration and can be expressed as the product of the derivative of the successive iterations by the chain rule of differentiation:

$$T_t^n = T(F^{n-1}(t)) \; \ldots \; T(F(t)) \; T(t))$$

There is the same number of Liapunov exponents as the dimension of the phase space of a dynamical system. The largest exponent is positive for chaotic attractors and non–positive for non–chaotic ones. There are two types of method for calculating Liapunov exponents, suggested by Wolf(1985) and Eckmann & Ruelle (1985) respectively. The second algorithm is usually preferred (see, for example, Conte and Dubois, 1988) and was used here. To implement this algorithm, model (3) is iterated. At each step the derivative is calculated. Then they are multiplied using the QR factorization to get

$$T_t^n = T(F^{n-1}(t)) \; \ldots \; T(F(t)) \; T(t)) = Q_n R_n \; \ldots \; R_1 \,,$$

where Q_n is orthogonal and R_n upper triangular with non–negative diagonal elements. Then the diagonal elements $(\nu_n)_{ii}$ of $R_n \; \ldots \; R_1$ lead to the exponents

$$\lambda_i = \lim_{n\to\infty} \frac{1}{n} \text{Log} \mid \text{diagonal elements } (\nu_n)_{ii} \mid,$$

where n is the number of iterations.

The algorithm was programmed in FORTRAN and was tested by calculating the exponents of a chaotic attractor of the Hénon map

$$x_{n+1} = 1 - a\, x_n^2$$
$$y_{n+1} = b\, x_n$$

There is a chaotic attractor when $a = 1.4$, $b = 0.3$; the Liapunov exponents for this attractor are known to be $\lambda_1 = 0.42$, $\lambda_2 = -1.6$ (Conte and Dubois, 1988). The first 2000 steps of calculation of the Liapunov exponents for this attractor are shown in Fig. 1; they converge quickly. The result after 20000 iterations was [0.4168 −1.6208].

Liapunov exponents for a chaotic attractor in (3) is calculated. Values of parameters used are: $\alpha=\gamma=1.0$, $\mu=8.0$, $\beta=3.25$. The plotting of this attractor can be found in Jarrett and Zhang, (1993). The first 5000 iterations are shown in Fig. 2; the convergence is apparent. The three exponents after 30000 iterations are [0.20 −0.02 −0.70]. Shown in Fig. 3a and Fig. 3b are Liapunov exponents calculated as a function of β, in company with the bifurcation diagram in Jarrett and Zhang, (1993). Values of parameters in this calculation are $\alpha=\gamma=1.0$, $\mu=7.0$. The bifurcation diagram is copied here in Fig. 3c for comparison. By comparing the diagrams it can be seen that the first exponents are negative for periodic attractors and are positive for chaotic ones.

Calculation of fractal dimensions

The dimension of an attractor is a lower bound of the number of state variables needed to describe a steady–state behaviour. It can quantify the complexity of an attractor. A strange attractor normally possesses non–integer dimension, called a *fractal dimension*. Although many types of fractal dimension have been defined in the literature, an algorithm due to Grassberger & Procaccia (1983) is employed here to estimate *correlation dimension*. The correlation dimension is defined based on the correlation function of an attractor. A correlation function is the average fraction of points within a certain radius r on the attractor. For a system like (3) the correlation function $C(r)$ is given by

$$C(r) = \lim_{N\to\infty} \frac{1}{N^2} \{ \textit{the number of points } (t(i),\, t(j)) \textit{ such that } |\, t(i) - t(j)\,| < r \} ,$$

where N is the number of points considered. The correlation dimension D_C is defined as

$$D_C = \lim_{r\to 0} \frac{\log\, C(r)}{\log\, r} .$$

That is, it is the slope of the plot of log(C(r)) versus log(r). There may be only a limited range of the graph which is straight with an approximately constant slope. Only this range of data is used to estimate the dimension. The model is first iterated and the transient removed to get the attractor. Then the correlation functions are calculated for different values of r. The log–log plot of $C(r)$ versus r is made by the MATLAB software and the slope is estimated by least squares, also in MATLAB. The program for this algorithm is tested by calculating the correlation dimension of the chaotic attractor in the Hénon map mentioned above. Fig. 4 is the log–log plotting from the calculation. The slope

or the correlation dimension estimated is 1.256, which agrees well to that in the literature, 1.261 (Grassberger and Procaccia (1983)).

The program is used for computing the dimension of the same chaotic attractor in the gravity model mentioned above. The log–log plot is shown in Fig. 5 and the dimension estimated is 1.8251, though the attractor lies in a three dimensional phase space.

3. THE LOGIT–BASED TRIP ASSIGNMENT MODEL

Another problem in network flow dynamics is an extension of the logit–based trip assignment model. The logit–based model was first developed by Dial(1971). It was assumed that an individual driver chooses alternative routes according to the route cost in the way modeled by the logit discrete choice model. The model can be written as

$$x_i^{rs} = v_{rs} \frac{\exp(-\theta c_i^{rs})}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs})}, \quad \forall i \in p_{rs}, \ \forall rs \in P,$$

$$c_i^{rs} = \Sigma_{l \in L} \delta_{li}^{rs} d_l, \quad \forall i \in p_{rs}, \ \forall rs \in P,$$

and

$$y_l = \Sigma_{rs \in P} \Sigma_{i \in p_{rs}} \delta_{li}^{rs} x_i^{rs}, \quad l \in L,$$

where:

P = set of origin–destination pairs of node,
p_{rs} = set of routes connecting origin r and destination s,
L = set of links comprising the network,
v_{rs} = total flow from origin r to destination s,
x_i^{rs} = flow on route i joining r and s,
c_i^{rs} = cost on route i joining r and s,
y_l = flow on link l,
d_l = cost on link l, and $d_l(y_l)$ is the link performance function,
$\delta_{li}^{rs} = \begin{cases} 1 & \text{if link } l \text{ is in route } i \text{ joining } r \text{ and } s \\ 0 & \text{otherwise.} \end{cases}$

This model is static and does not consider congestion effects. Iterative dynamics can be introduced by assuming that the allocation of flow to alternative routes at the current stage depends on the cost and so the flow of the previous stage, or,

$$x_i^{rs}(n+1) = G_i^{rs}(x_i^{rs}(n)) \tag{6}$$

$$= v_{rs} \frac{\exp(-\theta c_i^{rs}(n))}{\Sigma_{j \in p_{rs}} \exp(-\theta c_j^{rs}(n))},$$

where

$$c_i^{rs}(n) = \Sigma_{l \in L} \delta_{li}^{rs} d_l(y_l(n)) , \quad \forall i \in p_{rs}, \ \forall rs \in P ,$$

and

$$y_l(n) = \Sigma_{rs \in P} \Sigma_{i \in p_{rs}} \delta_{li}^{rs} x_i^{rs}(n) , \quad l \in L ,$$

The link performance function $d_l(y_l(n))$ relates the travel cost to the flow. It is normally assumed that it is a monotonically increasing function of flow; any feasible function can be used. See Braston(1976) for a review. Dendrinos and Sonis (1990) actually suggested the idea for this kind of extension, though to a more general individual discrete–choice model and without any further analysis.

Consider first the case where there is one O–D pair joined by two links, putting $v_{rs}=1$ and using single subscripts only, then (6) becomes

$$x_i(n+1) = G_i(x_i(n)) = \frac{\exp(-\theta c_i(n))}{\Sigma_j \exp(-\theta c_j(n))} , \tag{7}$$

where

$$c_i(n) = c_i(x_i(n)) .$$

It can be that (7) is similar to the one–dimensional gravity model with the exponential deterrence function, except that the cost function here takes a general form. Thus, the analysis of the one–dimensional gravity model in (2.1) can be applied to (7). It is easy to see that (7) is a decreasing mapping and so has a unique fixed point x_i^e. By setting $\mu=0$, $\beta=\theta$, and replacing t_i^e by x_i^e in (5) we have

$$\left| \frac{d}{dx_1} G_1(x^e) \right| = \theta x_1^e x_2^e \left(c_1'(x_1^e) + c_2'(x_2^e) \right)$$

$$\leq 0.25 \, \theta \left(c_1'(x_1^e) + c_2'(x_2^e) \right) .$$

So if

$$\theta \left(c_1'(x_1^e) + c_2'(x_2^e) \right) \leq 1/0.25 = 4$$

the fixed point is stable. Note that the cost function here is the link performance function. If the condition for stability is not satisfied as a result of changing of parameter(s) then the fixed point becomes unstable and a stable period–2 orbit emerges. The fixed point and the period–2 orbit are the only possible steady states that could occur in this model; any initial condition will be attracted to the stable fixed point or a stable period–2 orbit if the fixed point is nonstable.

The general model (6) has been experimented numerically for several simple road networks. Fixed point and period–2 orbits seem to be the main behaviour in the model. This outcome is not too surprising because it is consistent with the numerical study of the gravity model with the exponential deterrence function the two models are quite similar. The trip assignment model seems to be more complicated with links and/or routes overlapping. This difference, however, has not been found to make the behaviour of the model richer except causing large oscillations in link flows.

4. SUMMARY

The dynamical properties of the gravity model and the logit–based trip assignment model have been studied. Fixed points and period two orbits are identified in one–dimensional gravity models with the power and the exponential deterrence functions, and in a one–dimensional logit–based assignment model. The conditions for the stability of these orbits were found. Liapunov exponents and correlation dimensions are found to be positive and fractal for chaotic attractors found in higher–dimensional gravity models with the combined deterrence functions.

The study in this paper has shown that the behaviour of the dynamical models depends on the interaction of the flow and the cost on a link, or between an O–D pair. The mechanism is that trip makers make their choices of routes, origins, or destinations) based on travel costs; while the travel costs change with flows. An empirical study can be made to determine what the relations of the flows and the costs really are. It needs to be stressed that chaotic behaviour has been found in the gravity model with the combined deterrence function, which, according to (Ortúzar & Willumsen, 1990), can fit the observed data better than the other two forms of function.

The gravity model with the exponential and power deterrence functions, and the logit–based trip assignment model considered in this paper, share a similar form. The analysis of the stability of fixed points and the period–2 orbits in one–dimensional models needs be generalized to higher–dimensional models. Although in a higher–dimensional model the stabilities can be studied by bounding the eigenvalue of the derivative at a fixed point, the results are not very conclusive. Some other methods are being explored currently and it is hopped to report in the future.

Acknowledgements

The author thanks Mr. D Jarrett for helpful discussions. The research in this paper is made during the term of a research studentship of Middlesex University.

References

BRANSTON D (1976). Link capacity functions: a review. *Transportation Research*, **10**(4), 223–236

CONTE R & DUBOIS M (1988). Liapunov exponents of experimental systems. *Nonlinear Evolutions*, 767–779. Editor: LEON JJP, World Scientific.

DENDRINOS DS & SONIS M (1990). *Chaos and Social-Spatial Dynamics*. Springer-Verlag

DIAL RB (1971). A probabilistic multipath traffic assignment model which obviates path enumeration. *Transportation Research*, **5**(2), 83–111

ECKMANN J-P & RUELLE D (1985). Ergodic theory of chaos and strange attractors. *Reviews of Modern Physics*, **57**(3), 617–656

GRASSBERGER P & PRACACCIA I (1983). Measuring the strangeness of strange attractors. *Physica* 9*D*, **9*D***, 189–208

HOROWITZ JL (1984). The stability of stochastic equilibrium in a two-link transportation network. *Transportation Research*, **18B**, 13–28

JARRETT DF & ZHANG XY (1993). The dynamic behaviour of road traffic flow: stability or Chaos? *BCS Displays Group: Application of Fractals and Chaos*, 237–248. Editors: CRILLY AJ, EARNSHAW RA & JONES H, Springer-Verlag.

ORTUZAR JdeD & WILLUMSEN LG (1990). *Modelling Transport.* Wiley

SMITH MJ (1984). The stability of a dynamic model of traffic assignment—an application of a method of Liapunov. *Transportation Science*, **18**(3), 245–252

WOLF A, SWIFT JB, SWINNEY HL & VASTANO JA (1985). Determining Liapunov exponents from a time series. *Physica*, **16D**, 285–317

Figure 1. *Liapunov exponents for Hénon map*

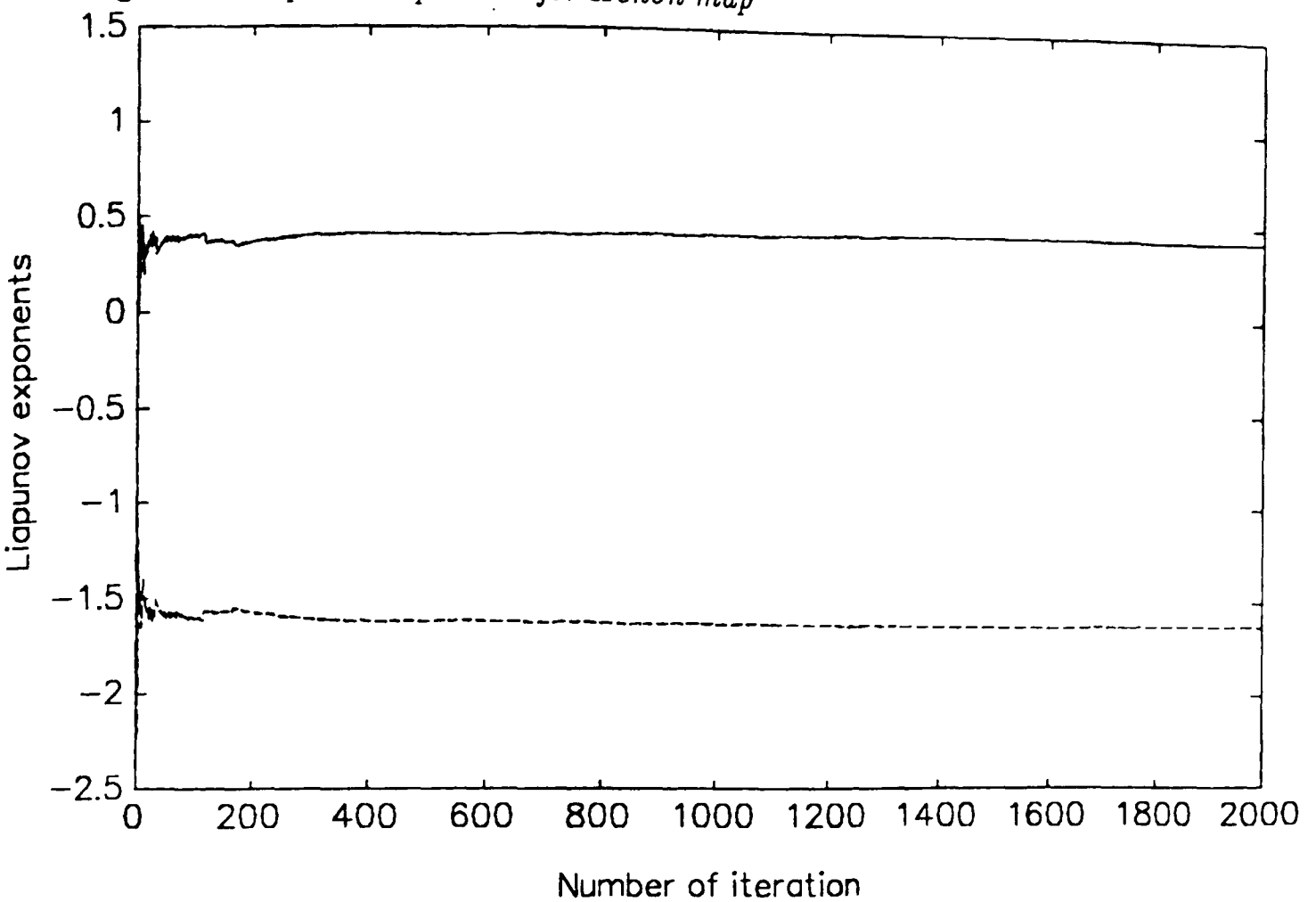


Figure 2. *Liapunov exponents for gravity model*

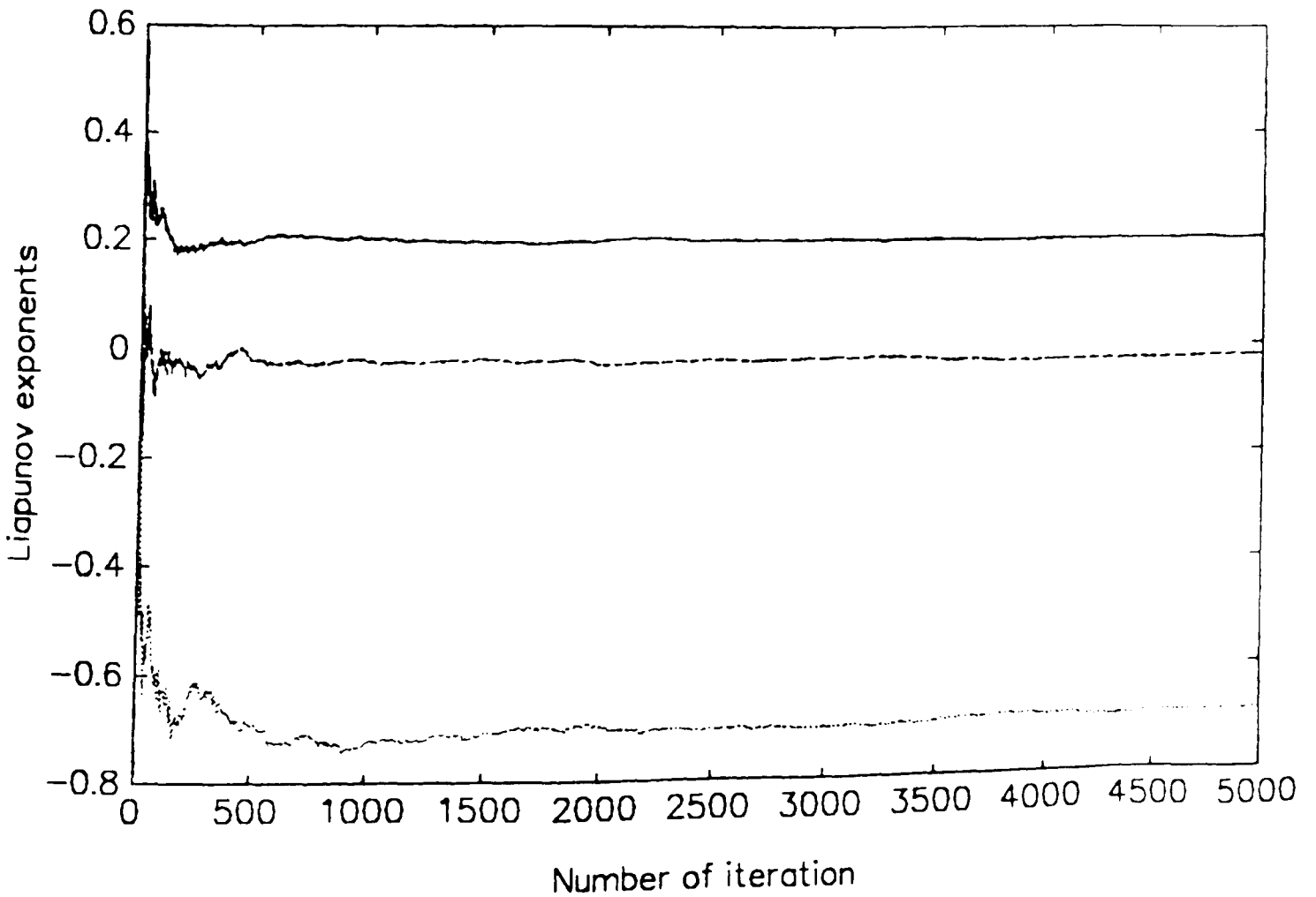


Figure 3a *The first Liapunov exponent for gravity model*

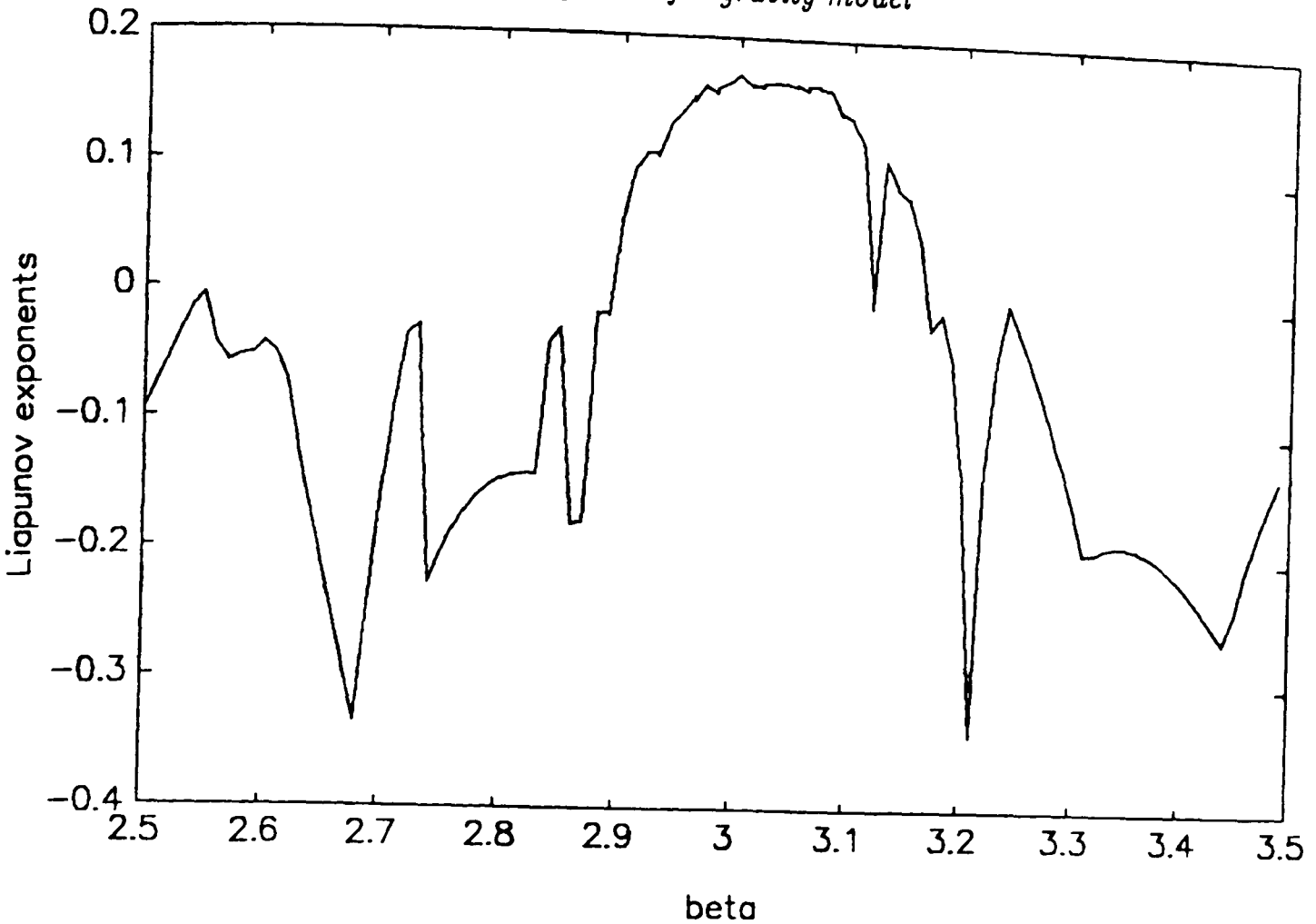


Figure 3b *The second and the third Liapunov exponents for gravity model*

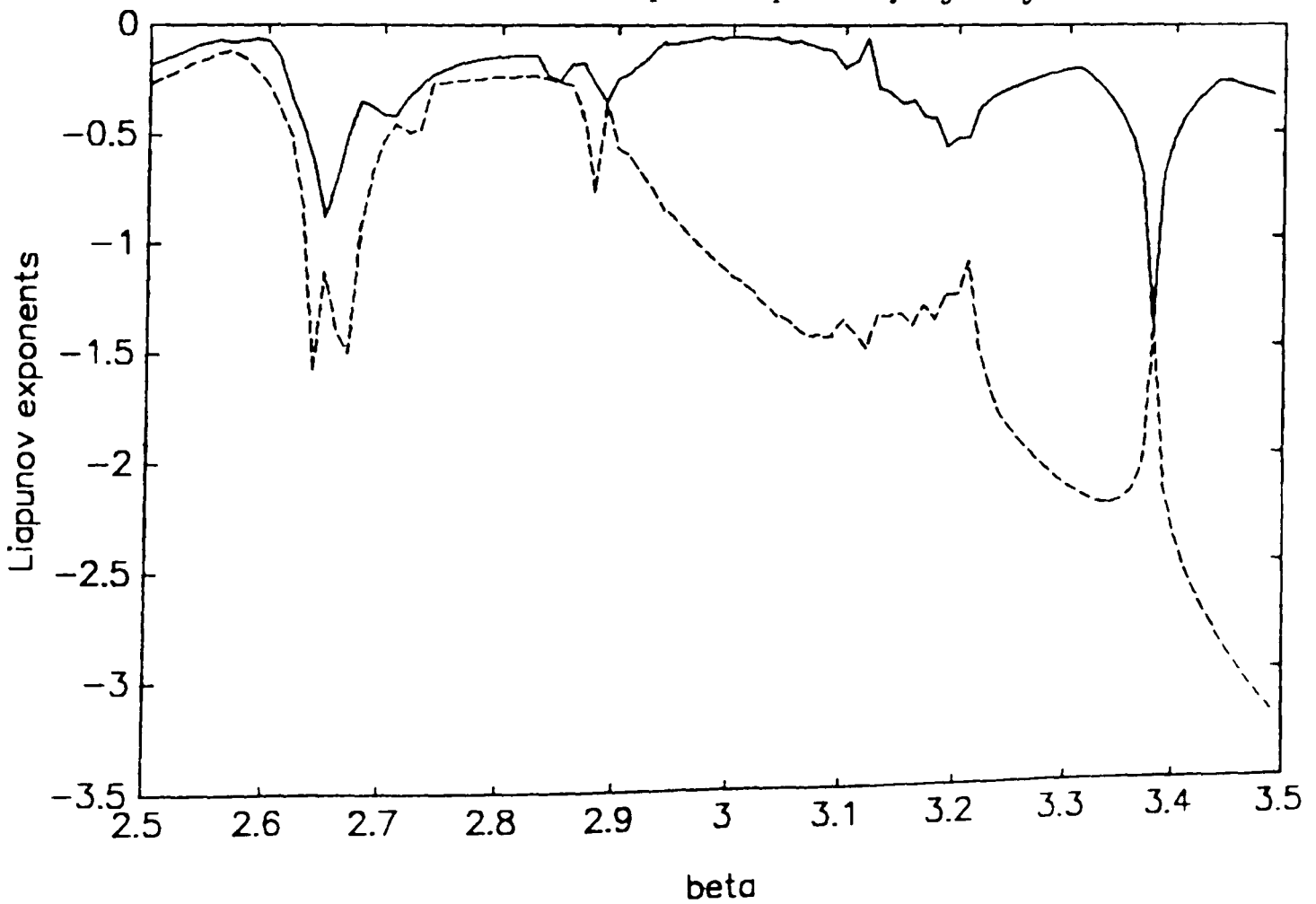


Figure 3c *Bifurcation diagram for gravity model*

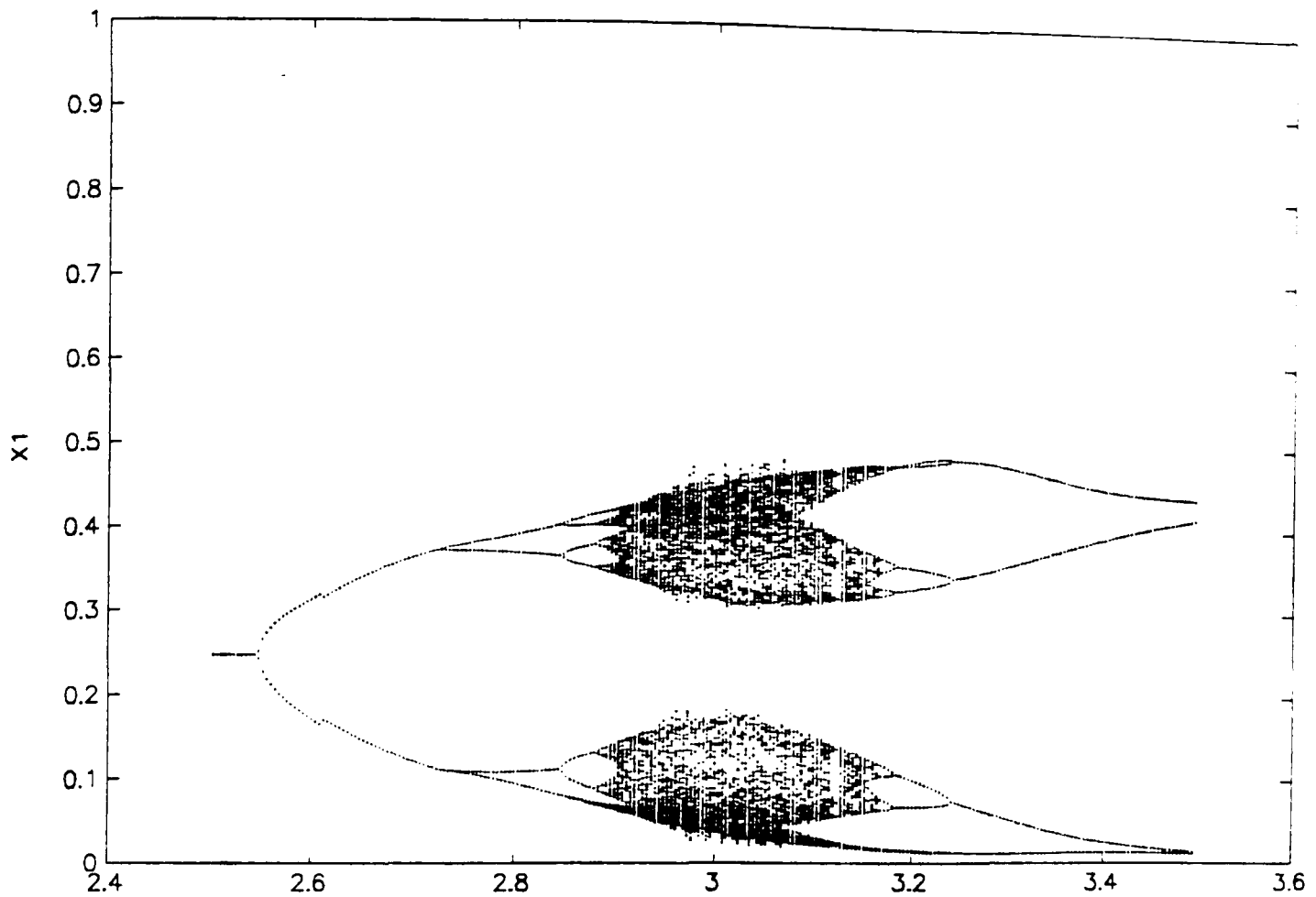


Figure 4. *Log* $(C(r))$ *versus* $log(r)$ *for Henon map*

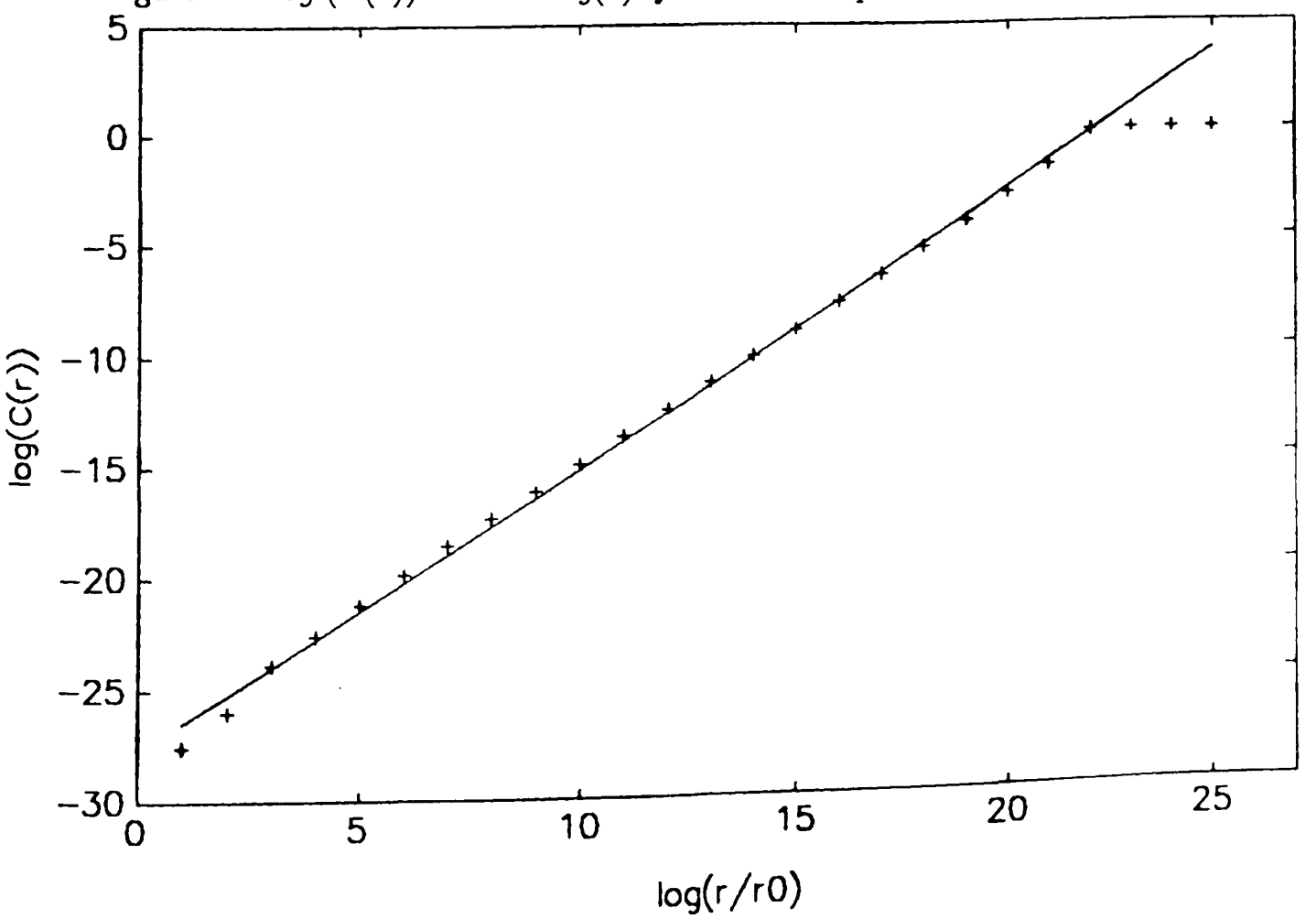
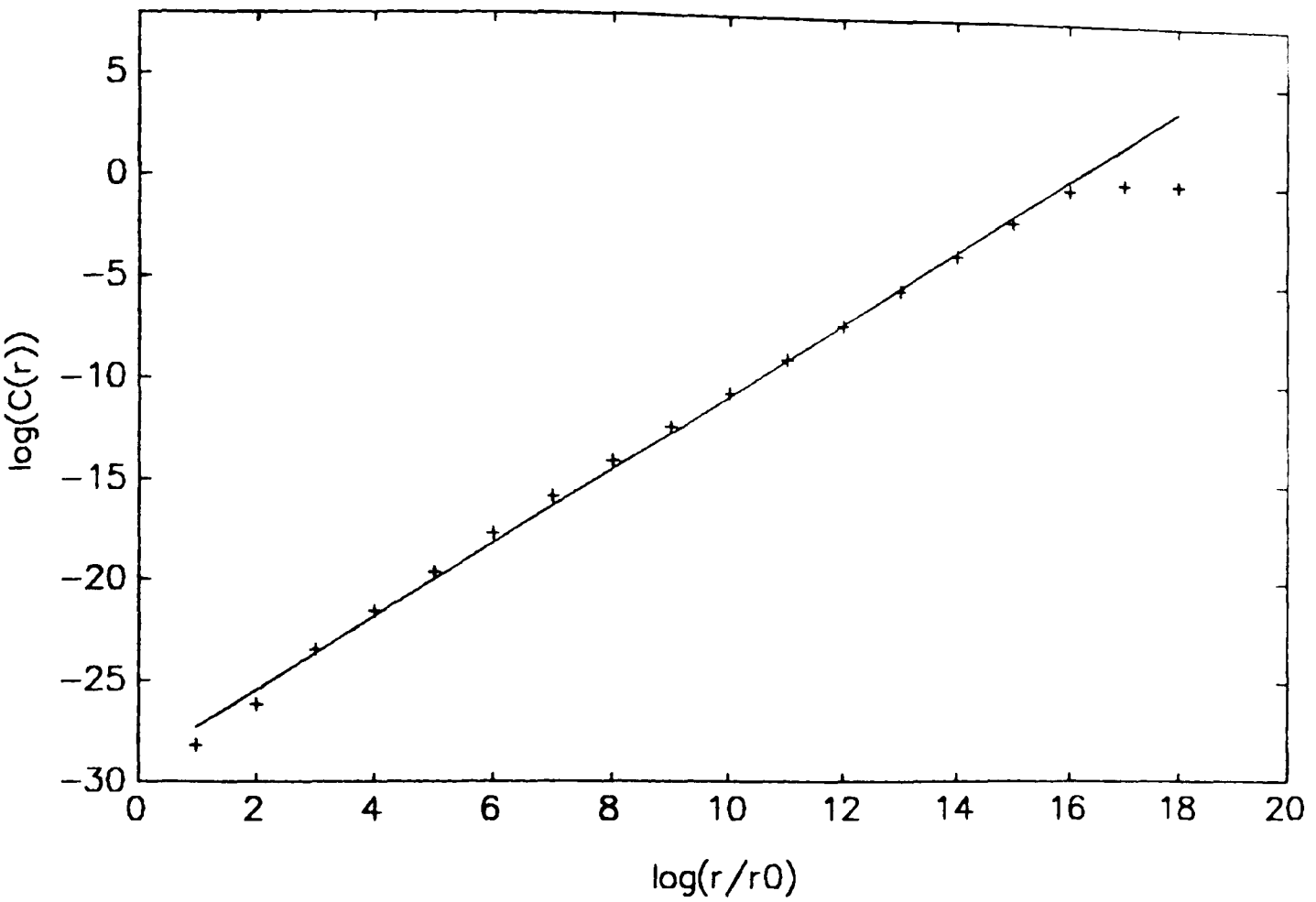


Figure 5. *Log (C(r)) versus log(r) for gravity model*



APPENDIX

This appendix gives the proves of the conclusion used in 2.1 about a one–dimensional map.

Given a map

$$F\colon \Delta_2 \to \Delta_2, \tag{A1}$$

where

$$F_i(\boldsymbol{x}) = \frac{g_i(x_i)}{\Sigma_j g_j(x_j)}, \quad g_i(x_i) > 0, \quad g_i'(x_i) < 0, \quad i = 1, 2,$$

then F has a unique fixed point, which bifurcates into a stable period–2 orbit as value of parameter increases. Trajectories from any initial conditions tend to either the stable fixed point, or a stable period–2 orbit if the fixed point is nonstable.

Prove

First of all, it can be shown that F_1 is a decreasing function of x_1 for

$$\frac{\mathrm{d}}{\mathrm{d}x_1} F_1(\boldsymbol{x}) = \frac{y_2 g_1'(x_1) + y_1 g_2'(x_2)}{\Sigma_j g_j(x_j)} < 0,$$

where

$$y_i = F_i(\boldsymbol{x}).$$

Let

$$P(x_1) = F_1(x_1) - x_1, \qquad x_1 \in [0,1]$$

and note that $F_1(0)>0$ and $F_1(1)<1$ because $F_1(\boldsymbol{x})$ is decreasing, then

$$P(0) = F_1(0) - 0 > 0,$$
$$P(1) = F_1(1) - 1 < 0,$$
$$P_{x_1}(x) = \frac{\mathrm{d}}{\mathrm{d}x_1} F_1(\boldsymbol{x}) - 1 < 0.$$

It follows that there is a unique point $x_1 \in [0,1]$ so that

$$P(x_1) = F_1(x_1) - x_1 = 0.$$

The point is a fixed point of (A1), denote this point by x^e, then

$$x_1^e = F_1(x_1^e)$$

The fixed point is stable if

$$\left| \frac{\mathrm{d}}{\mathrm{d}x_1} F_1(\boldsymbol{x}^e) \right| < 1 .$$

The derivative is always negative, so, if it is bigger than -1, disturbances decay oscillatory; while if it is less than -1, they grow oscillatory. In the latter case, the fixed point becomes unstable and a period–2 orbit appears through a flip bifurcation. To study the stability of the period–2 orbit it is convenient to consider the second iteration

$$G\colon \Delta_2 \to \Delta_2,$$

where

$$G_i(\boldsymbol{x}) = F_i(F(\boldsymbol{x})), \qquad i = 1, 2. \tag{A2}$$

It can be observed that G is a monotonically increasing map because

$$\frac{\mathrm{d}}{\mathrm{d}x_1} G_1(\boldsymbol{x}) = \frac{\mathrm{d}}{\mathrm{d}F_1} F_1(F) \frac{\mathrm{d}}{\mathrm{d}x_1} F_1(\boldsymbol{x}) > 0 .$$

A fixed point in G corresponds to a period–2 orbit in F. While the fixed point x^e in F is also a fixed point in G so that

$$x_1^e = G_1(x_1^e) .$$

Let

$$Q(x_1) = G_1(x_1) - x_1 , \qquad x_1 \in [0,1]$$

then

$$Q(x_1^e) = G_1(x_1^e) - x_1^e = 0.$$

and, if $\boldsymbol{x}^e$ is unstable,

$$Q_{x_1}(x_1^e) = \frac{\mathrm{d}}{\mathrm{d}x_1} G_1(\boldsymbol{x}^e) - 1 = (\frac{\mathrm{d}}{\mathrm{d}x_1} F_1(\boldsymbol{x}^e))^2 - 1 > 0 .$$

Therefore, in the neighborhood of $\boldsymbol{x}^e$, there are

$$\text{(a)} \quad Q(x_1) = G_1(x_1) - x_1 < 0 \quad \text{if } x_1 < x_1^e ,$$
$$\text{(b)} \quad Q(x_1) = G_1(x_1) - x_1 > 0 \quad \text{if } x_1 > x_1^e .$$

On the other hand, $G(\boldsymbol{x})$ is bounded on both sides and so is $Q(x_1)$ such that

$$\text{(c)} \quad Q(0) = G(0) - 0 \geq 0$$

$$\text{(d)} \quad Q(1) = G(1) - 1 \leq 0$$

In case of (a) and (c), there is at least one point $x_1^{p1} \in [\, 0, x_1^e \,)$ such that

$$Q(x_1^{p1}) = G(x_1^{p1}) - x_1^{p1} = 0$$

and

$$Q_{x_1}(x_1^{p1}) = \frac{\mathrm{d}}{\mathrm{d}x_1} G_1(\boldsymbol{x}_1^{p1}) - 1 = < 0 \Longrightarrow \frac{\mathrm{d}}{\mathrm{d}x_1} G_1(\boldsymbol{x}_1^{p1}) < 1$$

While in case of (b) and (d), there is at least one point $x_1^{p2} \in (\ \ 1]$ corresponding to x_1^{p1} so that

$$Q(x_1^{p2}) = G(x_1^{p2}) - x_1^{p2} = 0$$

and

$$Q_{x_1}(x_1^{p2}) = \frac{\mathrm{d}}{\mathrm{d}x_1} G_1(x_1^{p2}) - 1 = < 0 \Longrightarrow \frac{\mathrm{d}}{\mathrm{d}x_1} G_1(x_1^{p2}) < 1$$

The two fixed points is in fact a period–2 orbit in F and is stable because the derivatives lie between 0 and 1. The analysis here is made to a general model. Therefore, it can be concluded that stable fixed point and period–2 orbits are the only possible behaviour in a one–dimensional decreasing map. Conclusions proved.